NOVASYS

Novasys inc.
Bureau 2624
Tour de la Bourse
800, Place-Victoria
Case postale 151
Montréal (Québec) H4Z 1C3
Téléphone: (514) 875-7720
Télécopieur: (514) 874-9830

Université de Montréal

Université de Montréal
Faculté des arts et des sciences
Département d'informatique et de recherche opérationnelle
2920, chemin de la Tour
C.P. 6128, succursale Centre-ville
Montréal (Québec) H3C 3J7

Centre de recherche pour la défense de Valcartier

Exploration des techniques d'explication automatique, pour systèmes d'aide à la décision

Contrat no. W7701-8-1750/001/XSK

Autorité scientifique : Mme Micheline Bélanger

RECHERCHE ET DÉVELOPPEMENT POUR LA DÉFENSE

Défense nationale National Defence

Document final et prototype

Version 3.1

28 avril 1999

Defence Research and Development Canada | Recherche et développement pour la défense Canada

Canada

Français | Contact Us | Help | Search | Canada Site

Home | Defence R&D Canada | Defence Site

Browse Indexes

Titles

Authors

Descriptors

Corporate Authors

Corporate Acronyms

Report Numbers

Search Database

Summary

It is recommended you search "Authors" and "Corporate Authors" on their respective "Browse" pages.

Click here if you are having problems displaying text (ie. text is too big, or too small) or page appears to be jumbled.

Document: 1 of 1

Prior Top Bottom Next

Order

Click here to Order

Click here to Download

TITLE:(U) Exploration des Techniques d'Explication Automatique, pour Systemes d'aide a la Decision. Document final et prototype. Version 3.1

PERSONAL AUTHOR: Belanger, M.

CORPORATE AUTHOR: Defence Research Establishment Valcartier, Valcartier QUE (CAN);Novasys Inc, Montreal QUE (CAN);Montreal Univ, Montreal QUE (CAN) Departement d'Informatique et de Recherche Operationnelle

ABSTRACT: (U) L'une des lecons les plus importantes des dix dernieres annees de recherche dans le domaine de l'intelligence artificielle est que le rendement d'un systeme d'aide a la decision (SAD) n'en garantit pas necessairement l'acceptation par ses utilisateurs eventuels. Le manque de capacite pour ces systemes de justifier leur diagnostic constitue un handicap important a leur integration dans les milieux auxquels ils s'adressent. Le developpement d'une telle faculte devient donc crucial si l'on veut changer la perception actuelle des decideurs a l'endroit de tels systemes. C'est dans ce contexte que le Centre de Recherche et developpement pour la Defense de Valcartier (CRDV) a amorce une etude en collaboration avec les secteurs prive et universitaire sur les techniques d'explication automatique. Les objectifs de cette etude sont triples: 1. Faire une revue bibliographique de l'etat de l'art en matiere de techniques et systemes d'explication automatique; 2. Realiser une etude des avantages et des couts associes a l'ajout de mecanismes d'explication aux systemes d'aide a la decision; 3. Analyser les concpets d'explication automatique dans le contexte d'un SAD. TRUNCATED

DESCRIPTORS: (U) Artificial intelligence;Decision making;Bibliographies;Command and control;Expert computer systems;Problem solving;Knowledge bases;Software engineering;Reasoning;Adaptive systems;Learning theory;Benefit cost analysis

IDENTIFIERS: (U) Automatic explanation;Intelligence artificielle;SAD (Systeme d'aide a la decision);Commandement et controle

REPORT NUMBER: DREV-CR-1999-083; Contractor Report

REPORT DATE: 28 Apr 1999

NUMBER OF PAGES : 117

DRDKIM N°: 99-01414

CANDIS N°: 511364

Format(s): Hardcopy;Document Image stored on Optical Disk

Date modified: 2003-04-17

Important Notices

NOVASYS

Université de Montréal

TABLE DES MATIÈRES

HISTORIQUE DES VERSIONS DU DOCUMENT ..... **IV**

PRÉFACE ..... **VI**

ABRÉVIATIONS UTILISÉES ..... **VII**

1. INTRODUCTION **8**

2. QUELQUES RÉFLEXIONS SUR L'EXPLICATION **10**

2.1 À QUI EXPLIQUER? 13
2.2 À PARTIR DE QUELLES CONNAISSANCES? 13
2.3 LES DIVERS TYPES D'EXPLICATIONS 14
2.4 LES CONNAISSANCES POUR L'EXPLICATION : SOURCES ET EXPERTISE 16
2.5 L'EXPLICATION VUE COMME UN PROBLÈME À PART ENTIÈRE 17

3. DEUX TYPES D'ACQUISITION DES CONNAISSANCES EN IA **18**

3.1 AVANTAGES ET INCONVÉNIENTS DE CHAQUE MÉTHODOLOGIE 20
3.2 L'ACQUISITION DES CONNAISSANCES D'EXPLICATION 21

4. TECHNIQUES EXPLICATIVES **23**

4.1 PARAPHRASER LES RÈGLES (RÈGLES → LN) 23
4.2 ABSTRACTION PAR MÉTA-RÈGLES 24
4.2.1 Exemple de la stratégie de diagnostic de NEOMYCIN *25*
4.3 ENSEMBLE DE MÉTHODES ABSTRAITES 27
4.3.1 Classification Hiérarchique *28*
4.3.2 Mariage Structuré (Structured Matching) *30*
4.3.3 Conception de routine (Routine Design) *30*
4.3.4 Raisonnement fonctionnel *32*
4.4 MODÈLE CAUSAL 33
4.4.1 Exemple d'explications basées sur un modèle causal *35*
4.5 CBE (CASE BASED EXPLANATION) 36
4.5.1 Le projet SWALE *37*
4.6 MODÈLE EES (EXPLANABLE EXPERT SYSTEM) 38
4.6.1 La base de connaissances du modèle EES *39*
4.6.2 Module de construction automatique *41*
4.7 MODÈLE BASÉ SUR LA RECONSTRUCTION DES EXPLICATIONS 41
4.7.1 Le système REX *42*
4.8 LE RAISONNEMENT PROFOND 45
4.8.1 L'architecture du système CQFE *46*
4.9 DES EXPLICATIONS À L'AIDE D'AGENTS INTELLIGENTS 48

105-01824/

5. ANALYSE DES AVANTAGES ET COÛTS DES TECHNIQUES D'EXPLICATION 49

5.1 ANALYSE GÉNÉRALE DES AVANTAGES DES TECHNIQUES D'EXPLICATION 49
5.1.1 Qualité des recommandations 50
5.1.2 Aspects psychologiques 52
5.1.3 Processus de prise de décision 53
5.1.4 Analyse du comportement du système 54
5.1.5 Gestion des connaissances 55
5.2 ANALYSE GÉNÉRALE DES COÛTS DES TECHNIQUES D'EXPLICATION 56
5.2.1 Développement du système 56
5.2.2 Tests du système 58
5.2.3 Traitement du système 59
5.2.4 Aspects psychologiques 60
5.2.5 Gestion des connaissances 61
5.2.6 Environnement technologique 62
5.3 TABLEAU DE SYNTHÈSE DE L'ANALYSE DES AVANTAGES ET DES COÛTS 63

6. ANALYSE DÉTAILLÉE D'UNE ARCHITECTURE DE SYSTÈMES EXPLICATIFS 64

6.1 DIFFÉRENTS PROBLÈMES D'AIDE À LA DÉCISION 64
6.2 REPRÉSENTATION DES CONNAISSANCES 66
6.2.1 Différents niveaux de connaissance 66
6.2.2 Architecture proposée 67
6.2.3 Détails de l'architecture des bases de connaissances 70
6.2.4 Formalisme et représentation des connaissances 72
6.3 TRAITEMENTS SUR LES CONNAISSANCES 73
6.3.1 Introduction 73
6.3.2 Une architecture de traitement prometteuse basée sur une approche par agents intelligents . 74
6.4 INTERFACES 77
6.5 CONCLUSION 78

7. PROTOTYPE D'UN SYSTÈME EXPLICATIF 79

7.1 POINTS TECHNIQUES 79
7.2 MODE D'INSTALLATION 79
7.2.1 Installation de JDK1.1.7A 80
7.2.2 Installation de Swing 1.0.3 80
7.2.3 Installation du prototype 80
7.3 PROCESSUS GLOBAL DU PROTOTYPE 81
7.4 PRINCIPALES OPTIONS 82
7.4.1 Vue globale du système 82
7.4.2 Identification de l'usager 83
7.4.3 Mise en situation 84
7.4.4 Choix des poids et options 85
7.4.5 Décision et explication 86
7.5 PARAMÈTRES DU PROTOTYPE 87
7.5.1 Mise en situation 87
7.5.2 Paramètres à 0345Z 87
7.5.3 Options et actions (COA) associées 88
7.5.4 Explications selon les objectifs et les options 90

8. CONCLUSION 107

9. PERSONNES CONTACTS DU DOMAINE 108

10. BIBLIOGRAPHIE 110

HISTORIQUE DES VERSIONS DU DOCUMENT

Les modifications effectuées à la version 2 du document ont été identifiées par M^me^ Micheline Bélanger du CRDV.

Date	Version	Modifié par	Sections affectées	Raison
13/1/1999	1	Daniel Desjardins Esma Aimeur Simon Bélanger Claude Frasson	N/A	N/A
8/2/1999	2	Daniel Desjardins	Historique des versions du document	Ajout
8/2/1999	2	Daniel Desjardins	Préface	Modifications
8/2/1999	2	Simon Bélanger	Introduction	Modifications demandées
8/2/1999	2	Simon Bélanger	2.2	Modifications demandées
8/2/1999	2	Simon Bélanger	2.3	Modifications à l'introduction de la section
8/2/1999	2	Esma Aimeur	3.2	Modification demandées
8/2/1999	2	Esma Aimeur	3.3 et 3.4	Sections enlevées
8/2/1999	2	Simon Bélanger	4.3	Modifications demandées
8/2/1999	2	Simon Bélanger	4.3.1	Ajout demandé : classification hiérarchique
8/2/1999	2	Simon Bélanger	4.3.2	Ajout demandé : mariage Structuré
8/2/1999	2	Simon Bélanger	4.3.3	Ajout demandé : conception de routine
8/2/1999	2	Simon Bélanger	4.3.4	Ajout demandé raisonnement fonctionnel
8/2/1999	2	Simon Bélanger	4.4	Modifications demandées
8/2/1999	2	Simon Bélanger	4.5.1	Modifications demandées
8/2/1999	2	Simon Bélanger	4.7 et 4.7.1	Modifications demandées
8/2/1999	2	Simon Bélanger	4.8.1	Modifications demandées
8/2/1999	2	Simon Bélanger	4.9	Modifications demandées
8/2/1999	2	Daniel Desjardins	Analyse des avantages et coûts	Ajout correspondant à la phase 2 du projet
8/2/1999	2	Daniel Desjardins	Bibliographie	Modifications
24/2/1999	2.1	Daniel Desjardins	2.3	Modifications demandées
24/2/1999	2.1	Daniel Desjardins	3.2	Modifications demandées
24/2/1999	2.1	Daniel Desjardins	5.1.1.2	Modifications demandées
24/2/1999	2.1	Daniel Desjardins	5.1.2.1	Modifications demandées (anciennes sections 5.1.2.3, 5.1.2.4 et 5.1.2.6 regroupées avec 5.2.2.1)
24/2/1999	2.1	Daniel Desjardins	5.1.3.3	Modifications demandées
24/2/1999	2.1	Daniel Desjardins	5.2.2.1	Modifications demandées (ancienne section 5.2.2.2 annexée à 5.2.2.1)
24/2/1999	2.1	Daniel Desjardins	5.2.3.1	Modifications demandées
24/2/1999	2.1	Daniel Desjardins	5.2.3.3	Modifications demandées
24/2/1999	2.1	Daniel Desjardins	5.3	Modifications demandées

30 /3/1999	3	Daniel Desjardins	2.3	Ajout de textes et de références
30 /3/1999	3	Daniel Desjardins	5.1.4.1	Ajout de référence
30 /3/1999	3	Daniel Desjardins	5.1.4.2	Ajout de références
30 /3/1999	3	Claude Frasson, Esma Aimeur, Daniel Desjardins	6	Ajout d'un chapitre sur l'analyse détaillée
30 /3/1999	3	Simon Bélanger	7	Ajout d'un chapitre sur le prototype
30 /3/1999	3	Daniel Desjardins	Références bibliographique	Ajout d'un document contenant les références bibliographiques
28/4/1999	3.1	Daniel Desjardins	Page couverture	Ajouts : no. contrat et autorité scientifique
28/4/1999	3.1	Daniel Desjardins	6	Modification de structure du chapitre et ajout de précisions et de deux tableaux
28/4/1999	3.1	Daniel Desjardins	7.2.2 et 7.2.3	Ajout des adresses Internet pour Java 1.1.7a et Swing 1.0.3
28/4/1999	3.1	Daniel Desjardins	7.4	Ajout de précisions
28/4/1999	3.1	Daniel Desjardins	Conclusion	Modifications

PRÉFACE

L'une des leçons les plus importantes des dix dernières années de recherche dans le domaine de l'intelligence artificielle est que le rendement d'un système d'aide à la décision (SAD) n'en garantit pas nécessairement l'acceptation par ses utilisateurs éventuels. Le manque de capacité pour ces systèmes de justifier leur diagnostic constitue un handicap important à leur intégration dans les milieux auxquels ils s'adressent. Le développement d'une telle faculté devient donc crucial si l'on veut changer la perception actuelle des décideurs à l'endroit de tels systèmes.

C'est dans ce contexte que le Centre de Recherche et développement pour la Défense de Valcartier (CRDV) a amorcé une étude en collaboration avec les secteurs privé et universitaire sur les techniques d'explication automatique.

Les objectifs de cette étude sont triples :

1. Faire une revue bibliographique de l'état de l'art en matière de techniques et systèmes d'explication automatique;
2. Réaliser une étude des avantages et des coûts associés à l'ajout de mécanismes d'explication aux systèmes d'aide à la décision;
3. Analyser les concepts d'explication automatique dans le contexte d'un SAD.

Bien que ce travail de recherche aborde ces concepts dans un cadre général, le CRDV a pour objectif d'exploiter les résultats de cette étude pour développer des systèmes d'aide à la décision dans le cadre d'applications de commandement et de contrôle, notamment des systèmes d'évaluation de scénarios, de surveillance en temps réel, etc.

ABRÉVIATIONS UTILISÉES

AI	Intelligence Artificielle
BC	Base de Connaissances
CBE	Case Based Explanation
CH	Classification hiérarchique
CQFE	Ce Qu'il Fallait Expliquer
CR	Conception de routine
EES	Explanable Expert System
GT	Generic Task
MF	Modélisation fonctionnelle
REX	Reconstruction Explainer
RF	Raisonnement fonctionnel
SAD	Système d'Aide à la Décision
SBC	Système à Base de Connaissances
SBR	Système à Base de Règles
SE	Système Expert
SEE	Système Expert Explicatif
XP	eXplanation Patterns

1. INTRODUCTION

Le terme explication est utilisé dans la littérature des SAD pour faire référence à des domaines de recherche différents. Nous allons donc préciser ce que nous entendons par explication en nous servant du texte suivant, emprunté à R.C Schank [Schank, 82], qui met en évidence, deux des acceptions les plus courantes [Dominguez, 90].

> *Wish to claim that the fundamental difference between a system that can produce reasonable output and one that meets the criterion that is implicit in the term understanding system is that an understanding system should be able to explain its own actions.*
>
> *One might ask : explain its own actions to whom? There are two answers to this question, each of them quite important in their own way. The first answer is, of course those who inquire about how the system works. The second answer is, to itself. It is this latter answer that represents the essence of the understanding process, although, quite naturally, it is the lat««ter answer which is also the most difficult to assess.*

Si l'on s'intéresse aux systèmes capables d'expliquer leurs propres actions, plusieurs questions se posent : quelle est la finalité de ces explications? À qui s'adressent-elles? Quelle est la qualité de ces explications? L'idée est d'essayer de développer les capacités explicatives des systèmes, soit dans le but d'expliquer les résultats obtenus à un utilisateur externe (expliquer pour communiquer), soit dans le but d'améliorer le comportement du système lui-même (expliquer pour apprendre). Dans les deux cas, le système doit être capable de comprendre son raisonnement. Le problème de la compréhension est le même dans les deux cas; la différence principale réside dans les connaissances et les mécanismes que le système doit posséder pour pouvoir ensuite communiquer aux utilisateurs ses résultats.

Il faut distinguer ici les explications que l'on donne à un expert (et qui ne servent définitivement qu'aux experts capables de les comprendre) et les explications qui s'adaptent au niveau de l'utilisateur. Il est important de bien préciser cette différence car c'est justement à cause de la première approche (explications dédiées à l'expert) que beaucoup de systèmes n'ont pas su combler les attentes des utilisateurs. Il est donc primordial de chercher surtout à ce que le système donne des explications qui soient efficaces pour l'utilisateur. Cela veut dire qu'il les comprend, qu'elles sont utiles pour lui c'est-à-dire qu'elles l'aident dans son processus de prise de décision et à progresser et qu'elles sont adaptables selon les différents utilisateurs. De plus, sans cette approche, il est impossible de donner un certain niveau de confiance à l'utilisateur face aux recommandations données par les SAD.

Les travaux effectués dans ce domaine s'orientent vers l'amélioration des capacités explicatives des SAD et plus particulièrement des systèmes expert (SE)[1]. La plupart des

---

[1] Système ayant pour but la modélisation du comportement d'un expert humain accomplissant une tâche de résolution de problème, pour laquelle on ne dispose d'aucun algorithme, et ce dans un domaine précis. Il

SE utilisent en effet une approche « système à base de règles » (SBR). La supposition générale a été qu'une telle approche facilite la construction incrémentale de la base de connaissances (BC) et sa maintenance. Ainsi, les systèmes à base de connaissances (SBC) peuvent acquérir la capacité d'expliquer leur raisonnement. L'approche SBR présente toutefois des problèmes. Un de ces problèmes est que la structure de contrôle est rigide ce qui conduit les concepteurs des bases de connaissances à inclure dans les règles des informations de contrôle. Ces connaissances sont alors implicites et inaccessibles au système. Par conséquent, les capacités explicatives de ces systèmes sont limitées [Dominguez, 90].

Pour qu'un système d'aide à la décision engendre des explications plus complètes et compréhensibles, il doit être capable d'examiner le contenu de sa BC et son raisonnement. D'autre part, pour engendrer des explications, le système doit mettre en œuvre un raisonnement explicatif : il doit raisonner sur son propre raisonnement en vue d'engendrer des explications. Le raisonnement explicatif met en jeu deux types de connaissances. D'une part, toutes les connaissances qui se rattachent à la solution des problèmes : la BC du domaine d'application, les stratégies et les connaissances de contrôle, la trace de l'exécution des connaissances sur son propre fonctionnement, etc. D'autre part, les connaissances explicatives, c'est-à-dire, les connaissances permettant d'examiner les sources de connaissances pour engendrer des explications. Prenons l'exemple d'un fait que nous pouvons expliquer à l'aide de deux règles; une connaissance explicative pourrait nous recommander d'utiliser la règle qui contient le moins de prémisses, pour expliquer le fait en question. Expliquer est donc une tâche ardue nécessitant une importante quantité de connaissances. C'est pourquoi il est intéressant d'envisager une approche SBR pour réaliser cette tâche.

Cette étude sera amorcée à partir de différentes réflexions sur les concepts d'explication, entre autres, dans son application dans les systèmes experts. L'importance des connaissances explicatives et des méthodes pour leur acquisition sera présentée. Enfin, les principales techniques d'explication seront décrites et comparées sommairement en fonction de leurs caractéristiques spécifiques.

---

est composé essentiellement d'une base de connaissances, d'un moteur d'inférence et d'une base de faits (Lazure, 1988).

2. QUELQUES RÉFLEXIONS SUR L'EXPLICATION

L'explication dans les systèmes experts a été traditionnellement associée à deux types de questions : « comment » et « pourquoi ». Cette association provient du fait que ce sont les types de questions qui ont été mises en œuvre dans le module explicateur de MYCIN (un pionnier des SE), développé en 1974 par Davis [Shortliffe *et al*, 75]. Dans le cas général, la question « comment » consiste en une demande de justification de la présence d'un fait dans la BC. L'utilisateur peut demander comment une hypothèse a été formulée ou démontrée, comment un fait ou un diagnostic a été établi ou encore comment une caractéristique d'un objet a été déterminée. La question « pourquoi » par contre, porte sur le but que le système essaie d'atteindre. Une telle question est donc posée en cours de session au moment où le système demande à l'utilisateur des données qu'il ne peut déterminer lui-même à partir des connaissances dont il dispose. L'utilisateur demande alors pourquoi une donnée est demandée, quel est le plan de résolution adopté, pour quelle raison il est nécessaire de connaître un fait ou quelles sont les intentions du système.

Ces deux types de questions sont liés à des modes de raisonnement différents et leur interprétation en dépend. La question « comment » est surtout associée au raisonnement à partir des données *–chaînage-avant-*. Elle est interprétée comme étant la question suivante : quelles sont les prémisses de la règle qui a été appliquée pour déduire le fait en question. Ou plus généralement : quel est l'enchaînement des règles qui a conduit à la déduction du fait. La question « pourquoi », au contraire, est liée au raisonnement à partir des buts *–chaînage-arrière-*. Elle est interprétée de façon à répondre aux questions suivantes : quel est le but courant et quelle est la règle qu'on essaie de déclencher afin de le démontrer?

Un système raisonnant en chaînage-arrière peut facilement conserver une trace des déductions faites pour ensuite la présenter à l'utilisateur comme réponse à une question « comment ». Ainsi en partant d'un but, on arrive à comprendre comment les faits se sont enchaînés. Par contre, pour un système travaillant en chaînage-avant, il est difficile de répondre à la question « pourquoi » si c'est le système lui-même qui décide de l'enchaînement des règles sans aucune connaissance spécifique du domaine. En effet, en partant des données (chaînage-avant), on ne sait jamais vers quel but on se dirige. En définitive, dans certains systèmes, le raisonnement est dirigé par les données, mais n'est pas focalisé sur un but. Leur objectif est alors de saturer la BC et c'est pour atteindre cet objectif qu'ils posent des questions. D'autres systèmes se focalisent sur le but, mais le choix des règles est guidé par l'efficacité du processus plutôt que par une stratégie liée au domaine ou au problème à résoudre. Ces heuristiques « techniques » ne présentent pas en général un intérêt pour les utilisateurs qui veulent approfondir leurs connaissances du domaine d'application des SE.

Les questions « comment » et « pourquoi » sont très générales; leur interprétation correspond aux capacités d'un SBR de produire une trace de ses actions plutôt qu'aux

besoins des utilisateurs. Pour répondre à notre question initiale « que doit-on expliquer? », il faudra examiner d'abord les objectifs et intentions des utilisateurs, ensuite doter le système des connaissances nécessaires pour comprendre le raisonnement poursuivi pour résoudre un problème et répondre aux questions qui lui sont posées, et finalement réaliser un module capable d'exprimer les réponses dans un langage compréhensible et adapté aux utilisateurs.

Certains chercheurs ont effectué des exercices dans le but d'étudier les explications humaines ainsi que les types de questions auxquelles s'intéressent les utilisateurs des SE. Bien qu'il s'agisse d'expériences réalisées sur des domaines particuliers, nous pensons qu'il est intéressant de les examiner. Nous ne parlerons ici que de celles effectuées par R.W. Swartout [Swartout, 81] destinées à déterminer le besoin d'informations des utilisateurs et de celles menées par J.L. Weiner et son équipe [Goguen *et al*, 83] lors de leur étude des explications humaines. Nous mentionnerons ensuite les réflexions rapportées par W.J. Clancey dans [Clancey, 83] concernant notre question. Remarquons que ces trois recherches sont à l'origine de l'explication en tant que sujet de recherche à part entière et qu'elles ont, à notre avis, orienté la plupart des recherches postérieures.

Swartout a mené une expérimentation dans laquelle des étudiants en médecine devaient utiliser le système OWL de prescription de digitaline et poser (oralement) toute question qui leur venait à l'esprit. Les questions ont été enregistrées puis examinées dans le but de déterminer quelles sortes de questions un médecin peut se poser en utilisant ce système. Voici les trois types de questions répertoriées par cet auteur :

- Les méthodes que le système utilise (par exemple : « Comment avez-vous calculé la dose de digitaline? Elle est inférieure à celle que j'attendais »).

- La justification de ce que le système est en train de faire (ex : « Pourquoi voulez-vous réduire temporairement la dose de digitaline? »).

- La confusion sur la signification des termes utilisés (par exemple, à la suite de la question « la question fonction rénale est-elle stable? », un étudiant dit : « maintenant cette question ... je ne suis pas sûr ... fonction rénale stable », voulez-vous dire stable à l'état anormal ou bien ... c'est-à-dire que la fonction rénale du patient est anormale, mais elle est stable pour l'instant »).

Au regard de cette liste de questions, Swartout rappelle que le premier type de questions était déjà traité par le système OWL et que les efforts de sa thèse –conception et réalisation du système XPLAIN- se sont concentrés sur le deuxième type de questions; c'est-à-dire sur les questions portant sur la justification des actions du système.

Les travaux de Weiner et de son équipe s'orientent vers l'élaboration d'un modèle de la structure des explications humaines dont l'objectif final est d'être utilisé pour engendrer automatiquement des explications (pour les êtres humains). Ils s'appuient sur des transcriptions de conversations entre personnes pendant lesquelles les explications ont lieu de façon naturelle et non sur des protocoles obtenus en situations artificielles (qu'ils critiquent d'ailleurs). Les transcriptions examinées proviennent de trois sources : des

conversations entre l'ex-président américain Nixon et ses conseillers- Watergate transcription of Government Printing Office-, des interviews concernant le choix de carrière, et finalement des interviews concernant le choix de formulaire –long ou simplifié– pour la déclaration de revenus. L'analyse de ces transcriptions a été faite en vue de déterminer les différentes façons dont un sujet essaie de justifier une phrase. Ils y ont trouvé trois façons :

- Donner une raison de la forme « A because B » ou « B so A » (par exemple : « L'entreprise paie toutes mes dépenses, de sorte que « so that » - je n'ai pas beaucoup de dépenses d'affaires à reporter »);

- Fournir des phrases plus spécifiques –qu'ils appellent exemples-(ex : « Bien, je connais la plupart de mes propres déductions. Je sais quelles déductions sont avantageuses, comme la dépréciation de ma propriété, la maintenance de ma propriété, et des choses comme celles-ci, la déduction que j'obtiens sur les intérêts ... »);

- Éliminer graduellement les phrases alternatives, par exemple :

  1. « J'ai calculé mes déductions et elles ne ressemblaient pas aux déductions standards »;
  2. Transformation : « Je dis standard de telle façon que j'aurais probablement pu remplir un formulaire simplifié si je l'avais voulu »;
  3. Transformation : « C'est pourquoi j'ai rempli un formulaire simplifié si je l'avais voulu »;
  4. Transformation : « C'est pourquoi j'ai utilisé un formulaire long ».

Les réflexions de Clancey concernant le type de questions auxquelles un système à base de connaissances doit répondre sont issues de l'expérience acquise lors de la réalisation du système GUIDON. Le but principal du projet correspondant était de transformer le système expert MYCIN en un outil d'enseignement [Clancey, 87]. Il a constaté certaines limites des capacités explicatives de MYCIN, déjà mentionnées par Davis lors de l'implantation de la commande « pourquoi » dans ce système [Davis, 76], inhérentes à la conception initiale de la BC. En effet, la question « pourquoi » permet à l'utilisateur de MYCIN de connaître, en cours de session, la règle considérée et le but envisagé dans l'étape de raisonnement courant du système. Cependant, d'autres interprétations de cette question, concernant la justification des règles et la stratégie globale de résolution, ne peuvent pas être considérées. Par exemple, MYCIN ne peut pas répondre à la question « pourquoi examine-t-on l'hypothèse « méningite bactérienne » avant l'hypothèse « méningite virale »? Parce que même si le but à démontrer porte sur la première hypothèse, il considère toutes les règles qui concluent sur un type de méningite et non seulement celle qui permet de conclure sur la méningite bactérienne. Ces questions intéressent particulièrement Clancey, car l'objectif de ses travaux initiaux était d'enseigner l'expertise de MYCIN. La justification des règles facilite la compréhension et la mémorisation de ces règles et une vue d'ensemble de la stratégie pour ensuite l'utiliser dans la résolution de problèmes similaires.

L'aperçu des questions possibles, même en étant incomplet, montre que le nombre de questions qu'un utilisateur peut se poser est très important en ce qui concerne la connaissance d'un système, l'utilisation de cette connaissance dans la résolution d'un problème et les résultats obtenus.

2.1 À QUI EXPLIQUER?

Les explications portant sur le raisonnement d'un système expert peuvent être exploitées par un expert du domaine, un ingénieur de la connaissance, un étudiant ou un utilisateur quelconque. Chacun parmi eux a ses objectifs différents, son vocabulaire différent, ses attentes spécifiques au niveau de la stratégie d'explication (explications ponctuelles (données de temps en temps), graduelles (données étape par étape), globales (données à la fin), référencées (faisant référence à d'anciennes explications), avec exemples, etc.) ainsi que sa connaissance plus ou moins approfondie du domaine d'expertise du système et des mécanismes automatiques de résolution de problèmes. Ceci fait en sorte que leurs critères de confiance et d'acceptation des explications du SAD sont très spécifiques à cause du contexte d'utilisation et des responsabilités des utilisateurs, par exemple :

- Les experts et les ingénieurs les utilisent surtout comme un outil pour mettre au point la BC;
- Un étudiant, par contre, cherche à acquérir des connaissances du domaine et à développer ses capacités de résolution de problèmes pour pouvoir ensuite résoudre seul les problèmes;
- Le système peut également être utilisé par des utilisateurs qui ne connaissent que certains concepts du domaine et qui veulent approfondir leurs connaissances de ce domaine;
- Un autre cas de figure est celui où le système aide l'utilisateur à prendre des décisions rapides, en situation de stress et/ou avec des impacts majeurs en cas de mauvaises décisions;
- Un système expert peut aussi générer des explications qui seront sauvegardées en mémoire et utilisées a posteriori au besoin (exemple : système de contrôle ou de sécurité sans intervention humaine directe).

Dans tous les cas, le système doit donner une explication complète et concise pour que l'utilisateur puisse à la fois comprendre chaque étape du raisonnement et les connaissances mises en jeu pour l'effectuer et acquérir ainsi la conviction de l'exactitude du raisonnement mené.

2.2 À PARTIR DE QUELLES CONNAISSANCES?

Dans les schémas classiques des systèmes experts à base de règles, la BC contient la connaissance du domaine exprimée sous forme de faits et de règles (mémoire à long terme) tout comme les faits représentant les données du problème particulier à résoudre et les faits déduits pendant la résolution de ce problème (mémoire à court terme). Un module d'explication dispose donc de ces connaissances et dans la plupart des cas, il dispose aussi d'une trace du processus de déduction contenant au moins un pointeur entre les faits déduits et les règles appliquées pour les déduire. Ces connaissances permettent de rendre compte des déductions qui ont été faites suite à la reconstruction du chemin de déduction allant du fait sur lequel porte la question de l'utilisateur jusqu'aux faits initiaux. Pour ce faire, un module « explicateur » peut s'appuyer sur les connaissances déjà mentionnées, mais aussi sur certaines connaissances sur le fonctionnement du mécanisme d'inférences et sur le rôle des règles, de leurs prémisses et de leurs conclusions.

Une telle explication est toutefois très limitée car elle ne rend pas compte du comportement du système, des actions effectuées et de la stratégie suivie. Cela est dû principalement au fait que les systèmes à base de règles (SBR) classiques ne disposent pas de toutes les connaissances qui sont prises en compte durant le processus de déduction; ils ne peuvent pas en conséquence engendrer des explications donnant un aperçu général de ce processus. Afin d'améliorer les capacités explicatives des SBR, les systèmes doivent alors disposer non seulement des connaissances qui permettent de résoudre les problèmes par la production d'une trace des déductions effectuées, mais aussi, de méta-connaissances sur l'utilisation des connaissances, par exemple :

- Les stratégies de résolution de problèmes;
- Les connaissances permettant d'examiner et de comprendre le raisonnement;
- Les connaissances permettant de déterminer ce qui doit être présenté aux utilisateurs, comment le présenter et à quel moment.

Un SBR a alors besoin de plus de connaissances pour expliquer son raisonnement que pour résoudre un problème.

2.3 LES DIVERS TYPES D'EXPLICATIONS

La notion d'explication de *première génération* réfère au fait que les explications sont exclusivement générées à partir des traces de résolution de problèmes. Dans la *deuxième génération* de techniques d'explications, l'idée directrice est que l'amélioration des capacités explicatives nécessite une évolution de la représentation des connaissances vers l'explication des principes et des modèles qui relèvent du domaine d'application. Cette classification est évidemment incomplète car elle ne tient compte que d'un seul des aspects de l'explication : les sources de connaissances. Elle laisse de côté d'autres aspects également importants qui concernent d'une part, les utilisateurs auxquels sont destinées les explications et d'autre part, la structuration des explications et les interfaces homme-

machine. Les systèmes experts tels que NEOMYCIN et XPLAIN sont classés parmi les pioniers de la deuxième génération des systèmes d'explication.

Dans la présente section, nous allons nous intéresser aux divers types de questions auxquels un système expert explicatif (SEE) doit pouvoir répondre pour rendre compte des raisonnements effectués pendant la résolution des problèmes qui lui sont soumis.

Examinons d'abord les activités qu'un SEE réalise pour la résolution d'un problème. Étant donné qu'il s'agit d'un système à base de règles, ses principales activités sont de déterminer les règles applicables aux faits décrivant le problème posé et puis de les appliquer. En termes logiques, on dira plutôt qu'il effectue des déductions et en termes plus généraux, on dira qu'il obtient, soit des résultats intermédiaires, soit une solution. La question que l'on peut lui poser alors est comment tel résultat a-t-il été obtenu? Nous appelons *explication positive*, une explication engendrée pour répondre à ce type de question. Pour engendrer une explication positive, le système peut s'appuyer sur la connaissance codée dans la règle qui a été utilisée pour obtenir le résultat en question; il doit donc garder la trace des règles déclenchées. Il est donc clair que les questions portent en général sur un résultat qui n'a pas été obtenu à partir des faits initiaux par application d'une seule règle, mais d'un enchaînement de règles. Celui-ci peut être long et complexe. Pour répondre à la question, le système doit alors considérer le raisonnement en sa totalité et essayer de le condenser et de le simplifier afin d'engendrer des explications compréhensibles.

L'utilisateur peut également se demander pourquoi un fait qu'il attendait n'a pas été déduit. Il peut s'agir d'une méconnaissance ou d'un raisonnement erroné d'un utilisateur (*première hypothèse*) ou bien, d'une erreur dans la BC (*deuxième hypothèse*). Pour répondre à cette question, un système peut travailler à partir de la première hypothèse. Il essaie d'expliquer pourquoi le fait ne peut être déduit, autrement dit, pourquoi le raisonnement de l'utilisateur n'est pas correct. Nous appelons ce type d'explication : *explication négative*. Comme nous le verrons plus loin, la deuxième hypothèse relève plutôt de domaines tels que la validation des BC ou encore de l'acquisition de connaissances et non de l'explication. Le problème ne proviendrait donc pas de la qualité de l'explication, mais plutôt de la qualité de la base de connaissances elle-même.

Un SEE peut aussi être considéré comme un résolveur de problèmes. De ce point de vue, il doit réaliser un processus de raisonnement qui aboutira à une solution du problème posé. Ce processus peut être décrit globalement en termes de stratégies. Ces stratégies ont pour objectif de conduire les actions du système vers la solution du problème. Une question pourra porter alors sur la stratégie utilisée pour atteindre le but demandé. La réponse sera alors une explication sur la stratégie. Cette stratégie doit être définie selon des critères propres au domaine concerné et non selon les critères syntaxiques habituels des SBR (diminuer l'ensemble de règles candidates ou augmenter les gains d'information).

D'autre part, s'il s'agit d'un système qui travaille sur des informations incomplètes, il peut être amené en cours de session, à demander des informations complémentaires. L'utilisateur peut alors se poser des questions sur l'utilité de ces informations.

L'explication d'une question doit mettre en évidence l'utilité de l'information demandée dans la stratégie locale ou globale poursuivie pour atteindre le but.

Chandrasekaran et al [Chandrasekaran *et al*, 88] propose un autre type d'explication qu'il appelle : justification de la connaissance. Il s'agit en fait de questions qui ne sont pas associées à une situation particulière, mais qui portent sur une partie de la BC elle-même. Par exemple, la question : « Pourquoi une réduction des taxes est-elle une bonne idée pour diminuer le déficit commercial? » dont la réponse est la suivante : « La réduction de taxes encourage l'épargne, stimule l'investissement et augmente la production; ceci diminue les prix, augmente les exportations et réduit ainsi le déficit commercial ». Ces auteurs font remarquer que cette justification d'une partie de connaissances provient d'une compréhension profonde du domaine et qu'un tel raisonnement n'est pas nécessaire pour conclure que la réduction des taxes est appropriée si le système possède une règle qui associe la réduction de taxe à la diminution du déficit commercial.

Enfin, pour [Herrmann et al., 98], la génération des explications peut être basée sur deux approches différentes: statiques ou dynamiques. L'approche statique permet de conserver et de proposer à l'utilisateur des explications ayant un format fixe. Cette approche peu coûteuse simplifie les tâches de collection, de structuration et de génération des explications. Toutefois, la flexibilité de cette approche est très faible. L'approche dynamique offre beaucoup plus de flexibilité et la génération des explications est fonction du contexte et du profil des utilisateurs. Par contre, la complexité et les coûts de développement et de gestion de ces explications dynamiques sont beaucoup plus élevés.

2.4 LES CONNAISSANCES POUR L'EXPLICATION : SOURCES ET EXPERTISE

Un système explicatif doit être capable à la fois de résoudre des problèmes et d'expliquer son raisonnement. Pour atteindre ces objectifs, il doit disposer au moins de trois sortes de connaissances : en premier lieu, des connaissances du domaine d'application, en second lieu, des connaissances explicatives permettant de comprendre son raisonnement pour ensuite rendre compte des résultats obtenus et enfin, des connaissances de contrôle pour gérer les différentes étapes de la résolution du problème. Les connaissances de contrôle peuvent être soit indépendantes du domaine, soit spécifiques au domaine d'application.

Pour appliquer le système à un domaine particulier, l'expert doit fournir les connaissances nécessaires à la résolution d'un problème et au contrôle de cette résolution. La connaissance spécifique du domaine comporte un modèle conceptuel et un modèle du raisonnement. Le modèle conceptuel contient les définitions des concepts manipulés, et des propriétés et des relations qui caractérisent ces concepts. Le modèle du raisonnement comprend les connaissances permettant de déduire les informations à partir des informations existantes (règles) ainsi que des connaissances pour utiliser les connaissances, autrement dit, le savoir-faire sur la stratégie de résolution à adapter.

Le système, quant à lui, doit posséder des connaissances explicatives et des connaissances de contrôle indépendantes du domaine d'application. Les connaissances explicatives doivent permettre d'interpréter la question posée par l'utilisateur et d'y

répondre, c'est-à-dire, elles doivent permettre de déterminer le contenu et la structure de l'explication correspondante. Les connaissances de contrôle s'occupent de la gestion de tâches que le système doit accomplir pour résoudre les problèmes qui lui sont soumis. Ces connaissances sont nécessaires pour permettre d'étudier la faisabilité des projets suite au développement des maquettes qui peuvent être vérifiées avant de disposer d'une spécification complète des stratégies adaptées aux différents types de problèmes envisagés.

2.5 L'EXPLICATION VUE COMME UN PROBLÈME À PART ENTIÈRE

Un courant de recherche actuel envisage la production d'explications dans un système à base de connaissances (SBC) comme un problème à part entière. Il semble en effet pertinent de vouloir isoler cette fonction qui possède une certaine autonomie à la fois pour des raisons historiques (depuis le module d'explication générique de MYCIN, les fonctions de raisonnement et d'explication sont bien dissociées dans les systèmes experts) et pour des raisons de sens commun (le terme d'explication fait référence à une problématique relativement bien délimitée avec une sémantique claire). On parle alors de connaissances d'explication pour dénoter les connaissances nécessaires à la production d'une explication et de raisonnement d'explication pour caractériser le processus de production. Les connaissances d'explication peuvent être de natures différentes. On peut donner quelques exemples de connaissances d'explication [Lemaire *et al,* 96] :

- S'il existe un concept proche de celui à expliquer et connu de l'utilisateur, tenter une explication par analogie;
- Le nombre d'hypothèses faites à propos des connaissances de l'utilisateur doit être minimal;
- Une explication longue doit être divisée en sous-parties reliées par des méta-commentaires;
- Tel utilisateur préfère les explications graphiques.

Des questions de raisonnement, de contrôle, de validation et d'acquisition des connaissances se posent dès que l'on aborde ce problème de nature des connaissances explicatives dans une vision de SBC. Ce dernier point est crucial pour la résolution de cette problématique, à savoir quelle méthodologie est la plus adaptée pour l'acquisition des connaissances d'explications et, plus généralement, pour le développement d'un système d'explication. Après avoir analysé les deux modes extrêmes d'acquisition des connaissances en IA, nous étudierons ces méthodologies appliquées aux connaissances d'explication.

3. DEUX TYPES D'ACQUISITION DES CONNAISSANCES EN IA

L'opération d'acquisition de connaissances dans un système d'intelligence artificielle (IA) met classiquement en jeu, d'une part, un ensemble d'humains (informaticien, cogniticien, expert ou même utilisateur), et d'autre part, la machine. Selon le degré de participation de chacune de ces entités dans l'opération d'acquisition, on aboutit à deux méthodes extrêmes d'acquisition de la connaissance.

La première inscrit l'acquisition dans un cycle de développement déclaration - expérimentation dont la fréquence de parcours est élevée, et s'appuie largement sur l'outil informatique. C'est l'acquisition incrémentale, aussi appelée prototypage rapide. La seconde ne dénigre pas ce cycle, mais n'accorde qu'un faible rôle à la rétroaction, portant son effort sur l'étape de modélisation : modélisation - implantation. L'historique de ces deux approches est décrit dans [Krivine & David, 91].

La première approche a connu ses lettres de noblesse avec les débuts de l'IA et les premiers systèmes experts. Le formalisme des règles de production permettait de déclarer des éléments de connaissance et de les mettre en vrac dans une base de connaissances. L'uniformité et la simplicité du formalisme semblaient faciliter l'acquisition. MYCIN et PROSPECTOR en sont les meilleurs exemples; ils ont permis d'aboutir à des conclusions qui n'étaient pas prévues par leurs concepteurs. Et c'est là le revers de la médaille : le concepteur du système ne maîtrisait pas toujours le comportement de son système, surtout lorsque celui-ci résultait de la combinaison de coefficients numériques attachés aux règles. Des problèmes de mise au point se posaient lorsqu'on voulait représenter dans le système un comportement bien spécifié à l'avance puisque les connaissances de contrôle ne pouvaient être explicitées. Cela rendait en plus l'explication de ces dernières impossible.

Ces constatations ont conduit alors à un renversement complet des méthodologies d'acquisition. Puisqu'on ne pouvait facilement maintenir et expliquer un système à base de règles qui manquait de structuration, on allait rejeter l'acquisition incrémentale et porter l'effort sur l'opération préliminaire à l'utilisation de la machine : les spécifications. Selon cette vision, l'acquisition consiste alors à concevoir, indépendamment des formalismes informatiques, un modèle de la tâche à résoudre [Breuker & Wielinga, 89], [Krivine & Davis, 91]. La majeure partie de l'activité de modélisation a donc lieu hors de toute considération informatique et repose essentiellement sur les épaules du cogniticien. Cette tâche de modélisation se rapproche de la démarche scientifique d'explication du monde, comme en physique, en biologie ou en psychologie. On explique une situation, puis on code le modèle dans l'ordinateur; on comprend le problème à l'avance, puis on donne à la machine une représentation de cette compréhension.

Alors que la première approche nécessitait de la part de l'informaticien, d'imaginer des mécanismes complexes pour gérer le non-déterminisme[2] inhérent aux problèmes d'IA, la seconde déplace l'activité d'acquisition vers une tâche de cogniticien et de construction de modèles papiers du monde. La machine n'est plus utilisée comme acteur de l'acquisition, mais comme un support de mise en opération d'un modèle. De ce fait, les problèmes de non-déterminisme sont gérés essentiellement en amont lors de l'étape de modélisation préalable. Les raisonnements sont transparents et explicites et sont compréhensibles par les experts, mais les types de problèmes envisageables sont différents. Nous montrerons que l'acquisition par modélisation préalable exige que le problème soit formalisable par l'humain.

Une seconde distinction entre les deux approches a trait au type de comportement auquel on s'intéresse. L'acquisition par modélisation préalable essaie de modéliser un comportement existant, le plus souvent humain. Même si ses partisans affirment clairement que le résultat n'est qu'un modèle, c'est-à-dire une représentation simplifiée de la réalité, et que leur objectif n'est pas de reproduire le comportement réel, mais d'engendrer une représentation acceptable, il reste que ce modèle est issu d'une réalité qui existe hors du système informatique, et qu'il est essentiellement fondé sur celui-ci. Or, les capacités d'un ordinateur sont complètement différentes de celles d'un humain, en ce qui concerne la mémorisation, le calcul ou le raisonnement. Il peut donc être légitime de vouloir baser le comportement du système, non pas sur des capacités extérieures, mais sur les siennes. Si elle aboutit rapidement à une implantation exécutable, l'acquisition incrémentale permet d'étudier le comportement de la machine et de baser l'acquisition sur celui-ci.

Cette proposition entre deux modes d'acquisition des connaissances n'a pas pour objectif de privilégier l'un par rapport à l'autre. Il convient cependant de distinguer clairement les domaines d'utilisation de chacun, et de ne pas considérer l'un ou l'autre comme la solution applicable à tous les domaines; ils ont chacun leurs intérêts et limites respectifs et c'est là l'objet du paragraphe suivant.

[2] Nous utiliserons l'acception suivante du terme déterminisme : une procédure est dite déterministe lorsqu'elle est sans ambiguïté dans son déroulement. Le non-déterminisme caractérise alors les problèmes pour lesquels il n'existe pas de solution algorithmique.

3.1 AVANTAGES ET INCONVÉNIENTS DE CHAQUE MÉTHODOLOGIE

Du point de vue de l'explication, l'avantage de l'acquisition par modélisation préalable est évident : étant plus explicite, le raisonnement du SBC peut être plus facilement expliqué. Étant donné que notre sujet de préoccupation est l'explication et qu'une distinction classique est faite entre système expliqué et système d'explication, nous pouvons énoncer un premier principe :

Principe1 *: Vu les approches actuelles d'acquisition des connaissances et des méthodes actuelles d'explications sur le concept de transparence, le système expliqué devra être conçu selon la méthode d'acquisition par modélisation préalable.*

Ce principe limite donc le champ des systèmes expliqués aux problèmes qui pourront supporter une explication a priori. On rejoint là la distinction que fait [Kodratoff, 86] entre connaissance experte et fortuite : une connaissance est dite experte lorsqu'elle peut être expliquée, justifiée; à l'inverse, elle est dite fortuite. En d'autres termes, le système expliqué devra être composé de connaissances expertes.

Quels sont alors les avantages de l'approche incrémentale? Ils sont dans sa capacité à s'attaquer à des problèmes complexes. Ce terme possède une signification bien particulière. Selon Newell, un problème est dit complexe lorsqu'il est composé de nombreuses parties qui interagissent de manière non triviale [Nii, 86]. Plus récemment, les théories de la complexité précisent cette définition.

Un système est dit complexe si des interactions mutuelles se manifestent entre différents niveaux. Il faut dans cette conception, que se distingue une structure hiérarchisée (niveaux), mais dont l'articulation ne soit par linéaire [Levy-Leblond, 91].

Grâce à une approche progressive, l'acquisition incrémentale rend possible la confrontation à un problème dont la complexité interdit, de toutes les façons, une compréhension globale et donc une modélisation complète par avance.

3.2 L'ACQUISITION DES CONNAISSANCES D'EXPLICATION

Dans cette partie, nous allons instancier les points abordés précédemment sur le problème de l'acquisition des connaissances d'explication. Commençons par poser le second principe.

Principe 2 : *Le problème de la construction d'une explication est un problème complexe.*

Essayons de justifier cette affirmation par un exemple. Une étude du problème permet d'identifier quelques grandes tâches qui interviennent dans sa résolution. Ce sont, entre autres, la saisie de la question de l'utilisateur, l'interprétation de cette question, le choix d'une méthode d'explication appropriée, la construction de l'explication correspondante, la transmission de cette explication. Nous prétendons que ces tâches ne peuvent s'enchaîner de façon linéaire : le résultat de l'une dépend de manière évidente des tâches précédentes, mais aussi des tâches suivantes. Par exemple, la construction de l'explication est conditionnée par le choix de la méthode, mais c'est parfois seulement au moment de cette construction que l'on réalise que la méthode choisie est inadéquate, parce que, par exemple, l'explication résultante nécessitera des justifications qui l'alourdiront et la rendront inutilisable. On voit donc que le choix de la méthode dépend aussi de la phase de construction. Ceci ne peut être évité qu'à condition de prévoir relativement précisément le résultat de chaque méthode de manière à posséder suffisamment d'informations pour faire un choix déterministe. On se rend bien compte que cette approche est irréalisable dès lors que le nombre de méthodes potentielles est important. On saisit donc toute la complexité du problème. Il semble alors très difficile de pouvoir en extraire une structure déterministe a priori.

De plus, l'humain met en œuvre dans la construction de ses explications des mécanismes importants d'évaluation; il sait à peu près estimer la qualité d'une explication, son impact sur son interlocuteur et parvient à rattraper une explication qui était mal engagée. Ces compétences lui sont nécessaires en raison de la faiblesse de ses capacités de mémorisation; il lui est quasiment impossible d'envisager plusieurs types d'explication en même temps. Il lui faut donc faire un choix bien souvent définitif de la méthode d'explication, quitte à mettre en œuvre des mécanismes puissants de modification, voire de changement imperceptible de cette méthode. La machine possède les compétences inverses : ses capacités de mémorisation sont importantes; en revanche les mécanismes qui permettent d'évaluer une explication, de choisir rapidement la bonne méthode, d'en changer « en douceur » ne peuvent lui être pleinement inculqués au stade actuel des recherches. Il convient donc de baser l'acquisition des connaissances d'explication aussi sur les compétences de la machine et pas uniquement sur un comportement humain bien différent.

Enfin, le contexte d'interaction entre le système et l'utilisateur n'est pas connu tant que le système n'est pas opérationnel. Il est vain de vouloir tenter de modéliser complètement cette situation à l'avance. La modélisation ne serait alors basée que sur l'étude d'une interaction humaine (par exemple, entre un expert et un futur utilisateur du système) ou pseudo-humaine comme le permet la technique du magicien d'Oz, dont le mode de

communication et le déroulement diffèrent sensiblement. La complexité du problème de l'explication s'exprime donc par les points suivants :

- La situation du dialogue face à un écran ne peut être appréhendée tant qu'elle n'existe pas encore;
- La machine possède des modes de raisonnement et de communication qui n'ont pas d'équivalent dans le mode humain (navigation dans un hypertexte, par exemple);
- Pour nous, le discours de l'expert est la connaissance même à modéliser alors qu'il n'est que le véhicule de la connaissance en acquisition classique, ce qui rend difficile l'explicitation par l'expert de son comportement d'explication;
- C'est bien souvent l'utilisateur en situation qui sera seul capable de valider les explications, contrairement aux systèmes experts classiques pour lesquels l'acceptabilité d'une solution peut être déterminée au niveau des experts [Nii, 86].

L'intervention de l'utilisateur final dans la phase de modélisation préalable ne résoudrait pas le problème puisque le contexte qui supporte cette validation n'existe pas encore. Ceci justifie donc la présence d'une présentation informatique du système dans la phase d'acquisition.

Il ne s'agit pas là de rejeter l'analyse de corpus, mais de limiter sa portée. Elle est utile pour définir un premier comportement d'explication, pour initier la boucle d'acquisition, mais son intérêt s'arrête là. Elle ne peut, à elle seule, être l'unique source de l'acquisition. Nous pouvons maintenant énoncer le principe 3 :

Principe 3 : *Vu les approches actuelles d'acquisition des connaissances, la faible connaissance des mécanismes humains d'explication, la faible connaissance des mécanismes informatiques d'explication, le système d'explication devra être conçu selon la méthode d'acquisition incrémentale.*

L'acquisition incrémentale soulève donc deux problématiques : la constitution d'une représentation initiale des connaissances, (que l'on tâchera de réaliser le plus riche possible et qui servira de courroies de démarrage), et des mécanismes puissants de modification et d'adaptation de ces représentations.

4. TECHNIQUES EXPLICATIVES

Ce chapitre expose les principales techniques d'explication du raisonnement pouvant être intégrées dans un système d'aide à la décision (SAD). Il explique sommairement leurs origines, les concepts inhérents à ces techniques, les principales forces et faiblesses ainsi que différentes applications de celles-ci à partir d'exemples tirés de la littérature.

Ces techniques d'explication sont présentées de manière à refléter le mieux possible le niveau croissant de complexité, de complétude et d'efficacité de celles-ci dans un système d'aide à la décision.

4.1 PARAPHRASER LES RÈGLES (RÈGLES → LN)

Cette technique est directement reliée aux règles que le système expert utilise lors de la résolution d'un problème. Chacune de ces règles, constituant la trace d'exécution, est paraphrasée afin de générer une explication associée à chacune d'elle. L'explication ou la paraphrase est développée de façon à contenir la règle choisie et avoir la forme d'une phrase compréhensible, le plus souvent en langue naturelle. Afin que cette approche fonctionne, les termes utilisés pour la formation des règles doivent être connus et compris de l'utilisateur. De plus, les règles doivent suivre un raisonnement familier à celui-ci [Swartout & Moore, 93].

L'avantage majeur de cette méthode est sa simplicité. Une fois que le système expert fonctionne, le travail qui doit être accompli pour générer les explications est relativement minime en comparaison avec le travail derrière le moteur d'inférence (moteur de résolution). Cependant, la qualité des explications générées souffre de la représentation de bas niveau des règles présentes au cœur du système expert. En effet, chaque règle résout une petite partie du problème présenté. Il est donc difficile de visualiser l'interaction entre les règles ou la stratégie générale du système et la résolution d'un problème suite à l'isolation de chacune des règles utilisées. Ainsi, il est impossible de générer des explications sur la stratégie globale du système. L'utilisateur doit être en mesure de comprendre chaque explication reliée à chaque règle. De plus, un second problème s'impose lorsqu'une règle ne possède pas assez de connaissances pour générer certains types d'explications tels que la justification d'une action. Le raisonnement derrière le choix et le moment où une action doit être exécutée est réalisé par le concepteur du système expert lors de sa construction. Parce que ce raisonnement n'est pas enregistré, il ne peut être expliqué [Swartout & Moore, 93].

Le système MYCIN [Shortliffe, 76] est à l'origine des travaux en la matière, avec le concept de *trace de raisonnement*, composé essentiellement de la succession des règles de production appliquées par le système pour parvenir à un diagnostic. De telles explications sont assez bien adaptées à la mise au point d'un système. En revanche, comme mentionné ci-dessus, elles ne permettent pas réellement de comprendre un raisonnement car une telle compréhension nécessite de se placer à un niveau

d'abstraction plus élevé que celui des règles et de faire, en particulier, intervenir la stratégie de résolution [Haton *et al*, 91].

4.2 ABSTRACTION PAR MÉTA-RÈGLES

Une méta-règle est l'expression d'une méta-connaissance sous forme d'une règle de production « si alors ». Cette technique est donc basée sur les méta-connaissances. Les méta-connaissances sont des connaissances d'ordre supérieur qui permettent de choisir les connaissances pour accomplir une tâche donnée. Exprimer des connaissances est un problème difficile à résoudre. Comme pour tout problème, on a besoin de connaissances pour le résoudre; ce sont ici des connaissances pour exprimer des connaissances, et que, par conséquent, nous appellerons méta-connaissances [Pitrat, 96]. Elles correspondent au recul que l'on prend par rapport à un certain domaine d'activité et représentent souvent de ce fait une part notable de l'expertise humaine. Elles interviennent souvent, car elles sont liées à la façon d'utiliser un ensemble de connaissances ou stratégies du raisonnement ainsi qu'à l'acquisition de nouvelles connaissances [Haton *et al*, 91].

Dans la section précédente, nous avons démontré les limitations de la transformation d'une règle vers la paraphrase utilisée, en particulier, à la génération d'explications dans le projet MYCIN. Clancey a développé par la suite le système NEOMYCIN [Clancey, 86] avec l'intention d'éliminer les problèmes et limitations présents dans le système MYCIN. L'idéologie derrière NEOMYCIN est d'utiliser des méta-règles afin d'atteindre une certaine abstraction permettant de visualiser la stratégie globale derrière le raisonnement du système. Par l'analyse du système MYCIN, Clancey a noté que le système passait par plusieurs phases d'exécution distinctes telles que la prise d'informations sur le patient, l'établissement d'une hypothèse et le raffinement de cette hypothèse. Dans MYCIN, ces processus ou phases étaient implicites car la structure à plat des règles ne supporte pas leur représentation explicite [Swartout & Moore, 93]. Afin de rendre ces processus plus explicites, des modifications sur l'architecture supportant le système de règles doivent être effectuées. Dans un premier temps, toutes règles effectuant un processus associé à une tâche particulière doivent être regroupées à l'intérieur d'un même ensemble de règles.

Les méta-règles permettent de contrôler la propriété active/non-active des ensembles de règles. En plus d'activer les ensembles de règles, les méta-règles recommandent des domaines de règles spécifiques à la demande d'une requête particulière d'un utilisateur. Par la catégorisation des règles en ensemble de règles correspondant à des phases de processus et l'utilisation des méta-règles pour spécifier le contrôle entre ces ensembles, NEOMYCIN profite de règles beaucoup plus abstraites que les règles de bas niveau que l'on retrouve dans le système MYCIN. Par cette architecture appliquée au système de règles et à l'aide des méta-règles possédant une certaine abstraction, le système NEOMYCIN est en mesure de générer des explications sur la stratégie globale empruntée par le système. Sans cette représentation, il est difficile de donner des explications qui démontrent comment une action spécifique peut être attachée à la stratégie globale.

4.2.1 EXEMPLE DE LA STRATÉGIE DE DIAGNOSTIC DE NEOMYCIN

La stratégie de NEOMYCIN est explicite et indépendante du domaine. Elle est structurée en termes de tâches, voir figure 1. Une tâche représente un but ou un sous-but du méta-niveau. À chaque tâche est associée une liste ordonnée de méta-règles.

Les prémisses des méta-règles examinent la liste d'hypothèses, les derniers faits établis et les relations causales du domaine. Les actions des méta-règles permettent d'invoquer d'autres tâches et au plus bas niveau, d'appeler l'interpréteur des règles pour atteindre des buts ou pour appliquer des règles du domaine [Dominguez, 90].

Notons que l'intérêt de cette représentation de la stratégie, en ce qui concerne l'explication et l'enseignement, provient principalement de deux faits. D'une part, les méta-règles sont indépendantes du domaine : elles sont exprimées à l'aide des concepts de la connaissance de structure, lesquels permettent de faire référence aux connaissances spécifiques du domaine. D'autre part, la structure de tâches représente une méthode générale de résolution de problèmes de diagnostic. Le système d'explication de NEOMYCIN s'appuie sur ces connaissances pour engendrer des explications au niveau de la stratégie générale de diagnostic [Dominguez, 90].

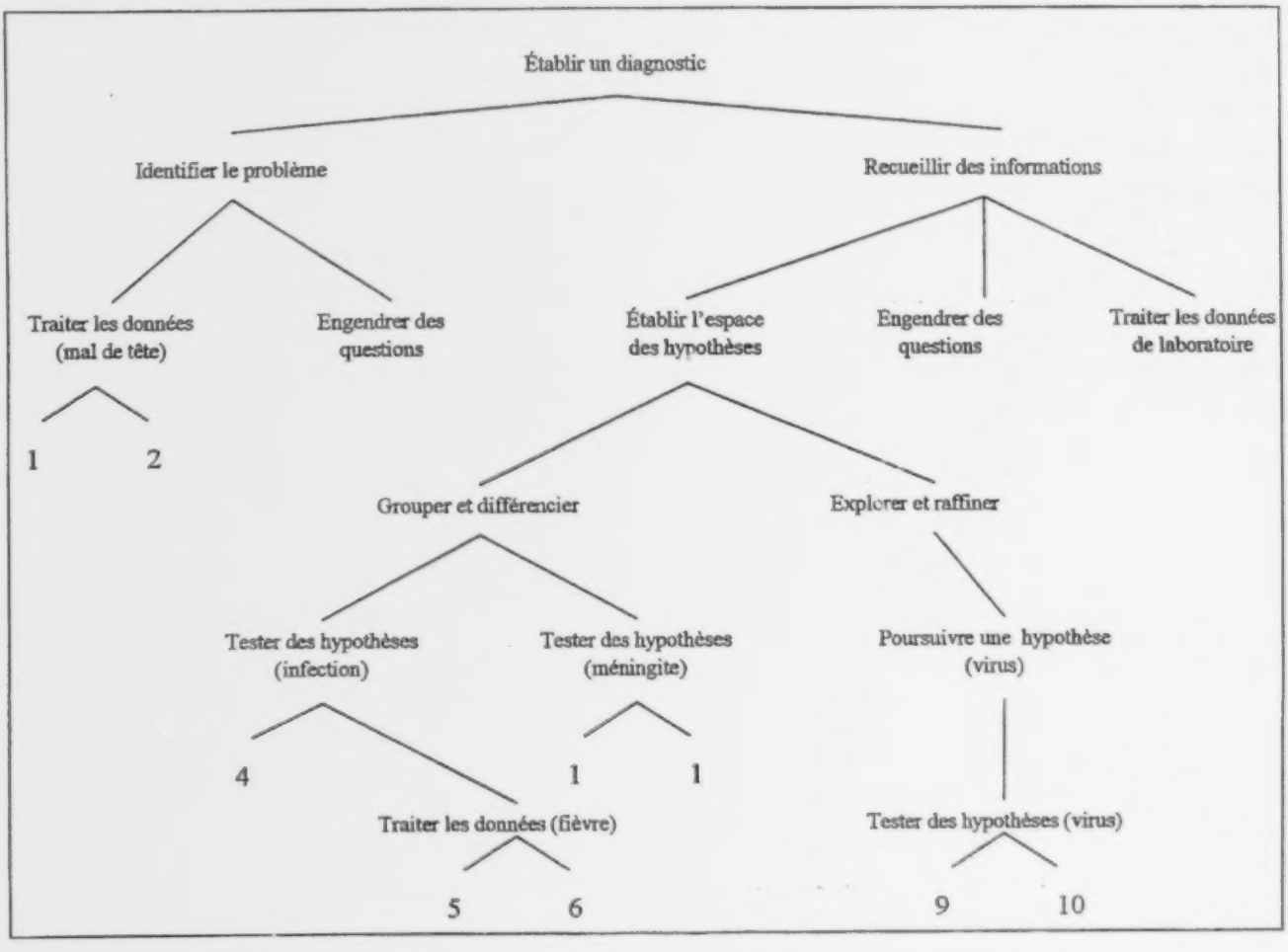

Figure 1 Représentation de tâches dans le système NEOMYCIN

Grâce à cette nouvelle représentation des connaissances, il est maintenant possible de poser la question *comment / pourquoi* en termes de tâches et méta-règles. Par exemple, dans l'extrait du dialogue présenté à la figure 2, la question 4) est une des questions posées par le système quand il est en train d'accomplir la tâche *tester une hypothèse*. La méta-règle qui a conduit le système à poser cette question est la méta-règle *411*. Cette dernière est utilisée pour répondre à une succession de questions *pourquoi;* le système remonte sur l'arborescence des tâches.

METAREGLE 411
SI il y a un fait fortement associé à l'hypothèse courante
ALORS appliquer les règles *déclic* concernées

EXTRAIT DE DIALOGUE-NEOMYCIN
4) Mary a t-elle de la fièvre?
** Pourquoi?
[i.e. pourquoi est-il important de déterminer si Mary a de la fièvre]

[4.0] On essaie de décider si Mary a une infection
La fièvre est fortement associée aux infections

Figure 2 Extrait du dialogue dans le système NEOMYCIN

La question *comment* concerne une tâche déjà effectuée. NEOMYCIN mentionne alors les méta-règles qui ont été appliquées et celles qui seront ensuite considérées, en vue d'accomplir la tâche en question. Tandis que les réponses aux questions *pourquoi* portent sur les prémisses de méta-règles c'est-à-dire, sur ce qui a permis d'appliquer la règle, les réponses aux questions *comment* portent sur les conclusions, c'est-à-dire sur ce qui a été fait.

4.3 ENSEMBLE DE MÉTHODES ABSTRAITES

Cette approche est basée sur l'observation que certaines stratégies (méthodes), ayant un niveau d'abstraction élevé, sont réutilisées à travers plusieurs systèmes experts afin de résoudre certaines classes de problèmes. Ces techniques sont caractérisées par leur démarche générale et la façon dont elles utilisent l'information disponible [Swartout & Moore, 93]. Chandrasekaran et ses collègues ont développé [Tanner *et al*, 93] un ensemble de méthodes abstraites, ayant comme nom « *Generic Task* » , bénéfique aux explications et à la construction de systèmes. Ainsi, il est possible, à partir de cet ensemble de méthodes, d'identifier des combinaisons tâche-méthode qui peuvent être utilisées dans la construction de solutions pour des tâches beaucoup plus complexes. Une tâche étant un problème ou un sous-problème que l'on soumet au système expert. Une méthode est combinée à une tâche lorsque l'on observe que cette méthode permet la réalisation d'une tâche, présente dans différents domaines, avec un taux de succès élevé.

Comme pour le système NEOMYCIN, l'amélioration des explications apportées par cette approche vient de l'abstraction des connaissances utilisées dans la résolution de problèmes. Cependant, dans NEOMYCIN, la structure abstraite de contrôle est régie par la structure des tâches et des méta-règles. Les explications sont donc produites à partir de

ces deux structures par un module d'explication plutôt général. Avec l'approche des GT, les méthodes abstraites ne sont pas représentées explicitement comme faisant partie du système. Les connaissances de ces méthodes sont seulement incorporées au module d'explication. Comme résultat, le module d'explication n'est pas généralisé, mais plutôt spécifiquement construit pour expliquer chaque type de tâche rencontré. Ainsi, les explications produites sont beaucoup plus compréhensibles pour l'utilisateur.

Les quatre prochaines sections présentent les principales méthodes abstraites. Elles sont traduites à partir d'articles du laboratoire des systèmes intelligents de l'université du Michigan.

4.3.1 CLASSIFICATION HIÉRARCHIQUE

La classification hiérarchique (CH) est de manière intuitive une organisation de connaissances et une technique de contrôle servant à faire une sélection parmi plusieurs options organisées de façon hiérarchique.

Le processus central de diagnostic peut être vu comme un processus de résolution de problèmes classificateur, i.e., qui identifie une description de cas comme un nœud dans une hiérarchie. L'approche est basée sur la décomposition des connaissances relatives au diagnostic en une hiérarchie de classification des spécialistes qui coopèrent. Ceci est fait de telle sorte que les spécialistes situés au haut de la hiérarchie représentent les hypothèses les plus générales, et que les sous-hypothèses de plus en plus spécifiques soient distribuées dans les niveaux inférieurs. Le régime de contrôle est un mécanisme d'établissement / raffinement de type «top-down» par lequel chaque spécialiste, lorsqu'il est invoqué, essaie de déterminer si les informations fournies par le cas présent supportent les hypothèses de diagnostic que le spécialiste représente (i.e. il essaie de se raffiner lui-même). En découpant l'espace d'hypothèse à de hauts niveaux de généralité, l'établissement / raffinement réduit la complexité de traitement inhérente au problème de diagnostic.

La figure 3 montre une décomposition fragmentaire de haut niveau de l'information pour le diagnostic d'une automobile en une hiérarchie de classification. Supposons qu'une voiture ayant un mauvais fonctionnement de sa batterie soit présentée au système, avec suffisamment d'informations pour porter un jugement sur le cas. Le contrôle procédera d'une manière «top-down»:

- Tout d'abord, «ProblèmeAuto» s'établit, puis se raffine. Durant l'étape de l'établissement, «ProblèmeAuto» utilise l'information disponible localement pour déterminer si la voiture est entièrement en mauvais état de marche. Durant l'étape de raffinement, «ProblèmeAuto» envoie à ses spécialistes subordonnés et plus spécialisés (SystèmeElectrique et SystèmeEssence) un message composé d'établissement / raffinement.

- Pour le cas soumis, le «SystèmeEssence» sera incapable de s'établir et par conséquent, aucun de ses subordonnés ne sera considéré pour le diagnostic.

- Contrairement au «SystèmeEssence», le spécialiste «SystèmeElectrique» s'établira et ultérieurement, se raffinera. Finalement, «Batterie» recevra un message d'établissement / raffinement et s'établira à son tour.

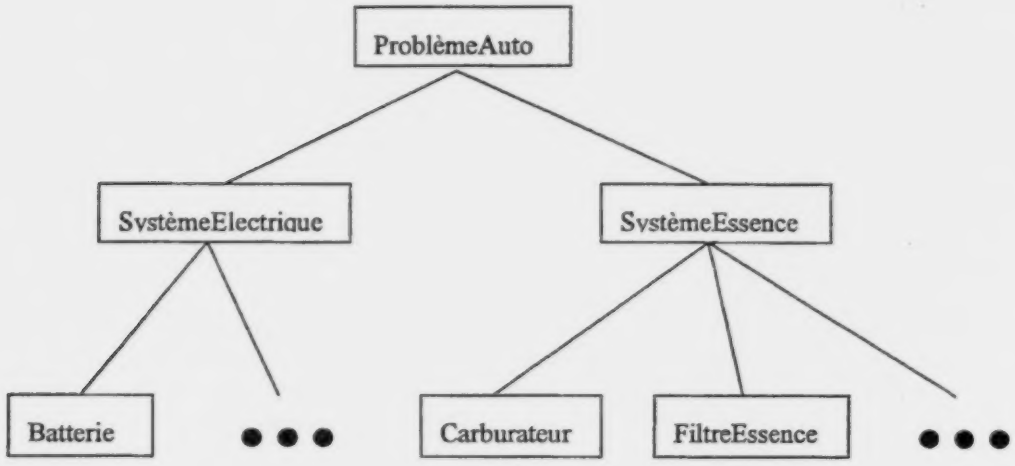

Figure 3 Décomposition fragmentaire de haut niveau de l'information

De cette façon, les spécialistes d'une hiérarchie de classification interagissent avec les autres spécialistes comme des composantes, envoyant et recevant des messages dans un vocabulaire fixe et à travers des chemins pré-déterminés.

4.3.2 MARIAGE STRUCTURÉ (STRUCTURED MATCHING)

Le mariage structuré est un mécanisme d'inférence simple pour réaliser des inférences de la forme: «Si les conditions A, B, C, ... sont vérifiées, alors la condition X est vérifiée». Les règles de cette forme sont organisées en ensembles structurés qui correspondent à une certaine structure de domaine.

La tâche du mariage structuré nécessite des abstractions symboliques hiérarchiques. Une abstraction de données est traitée sous forme de degré d'agencement. Ceci est symbolique, puisque l'abstraction est présentée comme l'une d'un petit nombre de mesures d'agencement qualitatives discrètes («fortement agencé», «faiblement agencé», «opposé», «fortement opposé», etc.), et hiérarchiques puisque l'abstraction finale est traitée à partir d'abstractions conceptuelles intermédiaires, qui à leur tour peuvent être traitées à partir d'autres abstractions intermédiaires, ou encore à partir de données brutes.

Étant donné que le nombre des valeurs symboliques possibles pour chaque niveau peut être gardé petit et qu'un certain nombre de combinaisons seront jugées impossibles par des experts, les tables qui incorporent de tels jugements sont typiquement clairsemées et faciles à traiter. Le processus est répété jusqu'à ce que les abstractions du plus haut niveau puissent être aisément traitées. Cette tâche de mariage (agencement) des hypothèses avec des données est une forme générale de raisonnement utile dans un grand nombre de contextes et est utilisée en conjonction avec les autres tâches génériques.

4.3.3 CONCEPTION DE ROUTINE (ROUTINE DESIGN)

La conception de routine (CR) a été proposée par Brown [Brown & Chandrasekaran, 89] comme une architecture pour faire de la conception et planifier des tâches pour lesquelles une expérience substantielle est disponible (et non pas pour la conception et la planification pour des situations totalement nouvelles). Brown a défini une méthode pour la conception de routine ainsi qu'un langage de représentation associé, DSPL (Design Specialists and Plans Languages) Cette méthode est appropriée pour les problèmes de conception du genre «routine». Il s'agit de situations de conception qui traitent de tâches répétitives rencontrées de façon routinière avec des besoins différents, mais similaires. Avec ce type de conception, le type de connaissances requises et les stratégies de résolution de problème sont tous deux connus à l'avance. La solution exacte, toutefois, n'est pas connue à l'avance. Être une « routine » ne rend pas un problème de conception simple puisque la résolution exacte du problème peut toujours être complexe et les possibilités peuvent être nombreuses, ce qui rend impossible la formulation du problème comme une recherche dans une table.

La conception de routine traite des situations de conception « de tous les jours » où les connaissances sur le « comment-faire » sont faciles à obtenir, comme par exemple la conception des composantes structurelles d'un nouvel édifice. Alors que le processus de

concevoir ces composantes peut être le même pour tous les édifices, le grand nombre de paramètres en entrée en cause rendent unique la solution à chaque cas.

L'approche de la conception de routine est basée sur une hiérarchie de spécialistes qui coopèrent, chacun étant responsable pour une partie spécifique et identifiée d'une conception complète. Les spécialistes de plus haut niveau dans la hiérarchie représentent typiquement les aspects plus conceptuels du processus de conception, alors que les spécialistes de plus bas niveau représentent les aspects plus paramétriques du processus de conception. Cette hiérarchie de spécialistes peut être vue, par analogie, comme un groupe de concepteurs humains qui coopèrent pour résoudre un problème de conception. Le spécialiste de niveau élevé peut être vu comme le chef de projet, avec ceux des niveaux intermédiaires comme chefs d'équipe. Les spécialistes de bas niveau dans la hiérarchie, quant à eux, sont comparables aux jeunes ingénieurs qui réalisent concrètement les traitements et les assignations de valeurs.

Dans la conception de routine, chaque spécialiste suit un plan parmi un ensemble préalablement spécifié de plans. Les plans prescrivent les actions à suivre pour la résolution de problèmes et sont définis au moment de la construction d'un système de conception. Alors que les parties du processus de conception (les plans) sont connues à l'avance, les combinaisons de ces parties ainsi que leur ordonnancement sont typiquement déterminés pendant la résolution concrète du problème. Et alors que chaque spécialiste possède un ensemble prédéfini de plans à partir duquel il peut choisir, le plan réel à suivre est choisi dynamiquement durant la résolution de problème, choix basé sur les besoins de conception fournis en entrée ainsi que sur les résultats de la conception établie jusqu'à là. Le choix des plans pour tous les spécialistes peut donc mener à la génération d'une conception nouvelle en raison de l'utilisation des plans par différents spécialistes qui n'ont jamais été utilisés ensemble auparavant. La conception de routine, toutefois, n'est pas appropriée pour les problèmes de conception complètement nouveaux.

4.3.4 RAISONNEMENT FONCTIONNEL

L'approche fonctionnelle à la représentation de mécanisme (RF) est née en 1983 de la recherche de Sembugamoorthy et Chandrasekaran ([Sembugamoorthy et Chandrasekaran, 84] in [Brown & Chandrasekaran, 89]). L'emphase de la recherche sur le RF a été mis sur le développement des techniques de simulation par l'utilisation des représentations fonctionnelles de mécanismes ([Sticklen, Chandrasekaran, & Bond, 89] in [Brown & Chandrasekaran, 89]). Le résultat est l'approche combinée représentation / raisonnement appelé la «modélisation fonctionnelle» (MF). Elle se distingue des autres recherches motivées par le RF.

Les intuitions sous-jacentes de la MF sont en trois volets. Premièrement, connaître les objectifs d'un mécanisme permet d'organiser la compréhension causale de son «fonctionnement». Deuxièmement, en MF, la causalité est représentée par de grands modules qui sont indexés par l'objectif des mécanismes. Ceci contraste avec plusieurs approches de réseaux de causalité où le mécanisme est représenté de façon monolithique par un réseau causal à un seul niveau. Finalement, le comportement global d'un mécanisme dans une situation donnée peut être compris par la composition de fragments appropriés d'un réseau causal.

Une représentation fonctionnelle utilise quatre composantes de bases: des mécanismes, des fonctions, des comportements, et des variables d'état. Les mécanismes sont des composantes actives du système qui est représenté et sont généralement représentées par une hiérarchie de composantes (voir figure 4).

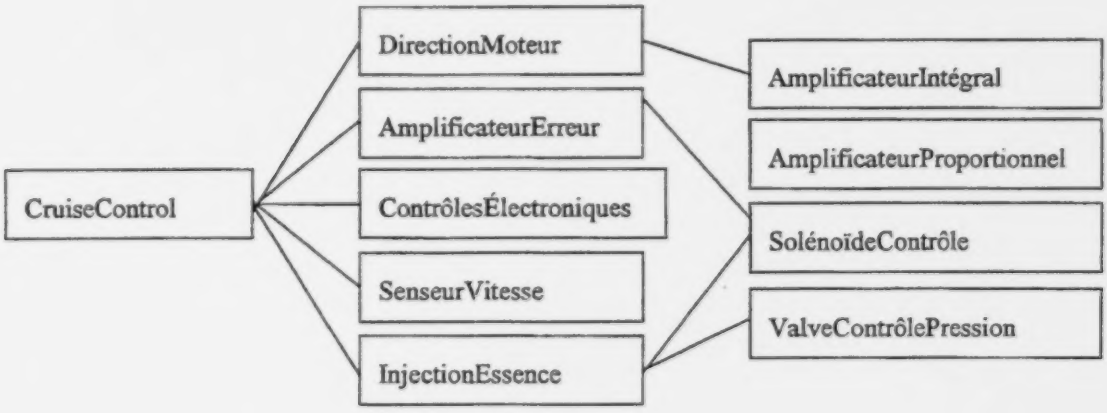

Figure 4. Exemple d'une hiérarchie de mécanismes

Les fonctions représentent les capacités des mécanismes, déclarées de manière abstraite. Les comportements sont les implantations des fonctions; ils décrivent comment les

fonctions sont réalisées en montrant explicitement les changements d'états. Finalement, les variables d'état décrivent les aspects intéressants du système.

4.4 MODÈLE CAUSAL

Cette technique implique un modèle causal associé à la partie active d'une base de connaissances visant les explications. Le modèle causal a pour rôle de contenir l'information décrivant les différents objets du système de même que les relations entre eux [Rolston, 88]. La structure d'un modèle causal est ordinairement un réseau sémantique tel que présenté à la figure 5.

Un réseau sémantique est un graphe étiqueté. Dans les nœuds figurent des objets ou des concepts et dans les arcs étiquetés, des relations entre ces concepts. Issus des travaux de psychologie cognitive sur l'organisation de la mémoire (mécanisme d'association), les réseaux sémantiques ont donné lieu à de nombreux systèmes et langages de représentation de connaissances tels que NETL, KL-ONE, etc. [Findler, 79]. Un réseau permet notamment d'exprimer une hiérarchie ou une taxonomie de concepts et la proximité entre des connaissances.

Les réseaux sémantiques constituent donc un moyen de structuration des bases de connaissances, problème important lorsque celles-ci sont de grande taille. Grâce au mécanisme d'héritage de propriétés, un nœud d'un réseau sémantique peut hériter des propriétés des nœuds qui le subsument. Ce mécanisme constitue un mode économique de raisonnement, permettant notamment à un concept particulier de posséder automatiquement les propriétés des concepts plus généraux dont il est issu [Haton *et al*, 91].

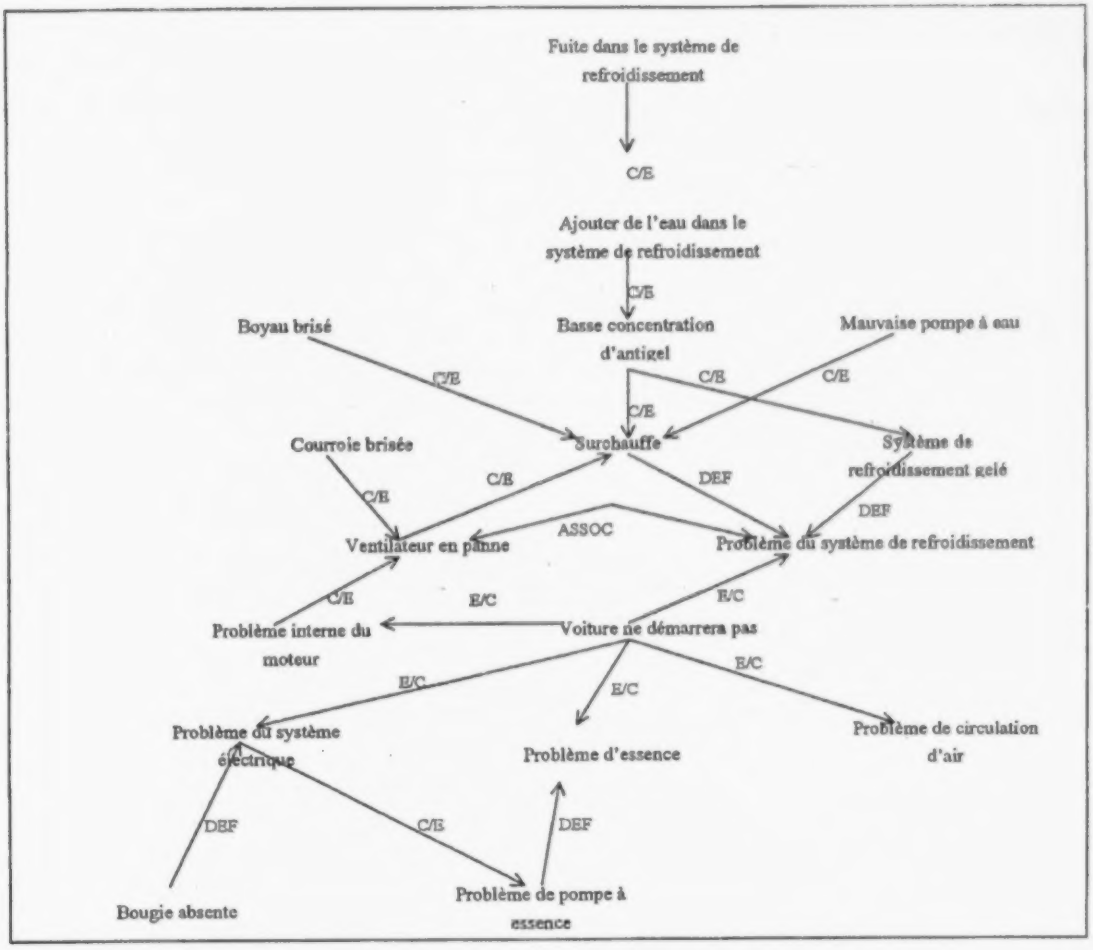

Figure 5. Exemple d'un modèle causal

Plusieurs types de relations existent sous ces modèles :

- Cause – effet : La conclusion qui suit les conditions préalables.
  Exemple :
  SI la jauge à essence est fonctionnelle et
  la jauge montre qu'il y a de l'essence,
  ALORS Il y a de l'essence dans le réservoir.

- Effet – cause : La cause peut être inférée par la présence d'un effet.
  Exemple :
  SI Un lieu est noirci,
  il y a une forte odeur de fumée et
  de la cendre est présente sur le lieu.
  ALORS Il y a sûrement un feu à cet endroit.

- Associative : Les conditions préalables et conclusions sont reliées, mais il n'y a pas de sens au niveau de la relation.
  Exemple :
  L'action d'augmenter les salaires est associée à l'inflation. Il faut noter que l'on ne peut pas déterminer si c'est l'inflation qui fait augmenter les salaires ou l'inverse.

- Définitionnelle : La conclusion décrit davantage la condition préalable, ce qui identifie fréquemment un objet comme un membre d'une classe.

4.4.1 EXEMPLE D'EXPLICATIONS BASÉES SUR UN MODÈLE CAUSAL

Les explications sont produites de façon dynamique par le système lorsqu'il traverse le modèle causal dans la résolution de problèmes. L'exemple ci-dessous présente comment les options *Pourquoi* et *Comment* peuvent être interprétées par un système expert hypothétique basé sur le modèle causal de la figure 5.

Utilisateur demande : Comment une fuite dans le système de refroidissement peut empêcher la voiture de démarrer?

Le système traverse les liens du modèle causal et répond :

Fuite dans le système de refroidissement peut causer
Ajouter de l'eau dans le système de refroidissement, qui peut causer
Basse concentration d'antigel, qui peut causer
surchauffe, qui peut causer
Voiture ne démarrera pas

Ou

Fuite dans le système de refroidissement peut causer
Ajouter de l'eau dans le système de refroidissement, qui peut causer
Basse concentration d'antigel, qui peut causer
Système de refroidissement gelé, qui peut causer
Voiture ne démarrera pas

Dans le cas où le système aurait déjà établi qu'il y a un problème avec le système de refroidissement et que l'utilisateur demande « Pourquoi *voiture ne démarrera pas?* », le système donnera comme explication :

Voiture ne démarrera pas est causé par
Problème du système de refroidissement, qui est associé avec
Problème interne du moteur, qui inclut
Problème de surchauffe, et
Système de refroidissement gelé.

4.5 CBE (CASE BASED EXPLANATION)

Il est impossible d'anticiper toutes les circonstances qu'un programme peut rencontrer lors d'une exécution. Même si cela était possible et que nous avions le pouvoir de construire toutes les réponses appropriées, le monde en constante évolution est un obstacle majeur aux bases de connaissances statiques. Un module d'apprentissage devient donc essentiel aux systèmes qui interagissent avec la réalité [Schank & Leake, 89].

Lorsqu'une personne fait face à une nouvelle situation, il est peu probable qu'elle essaie de résonner à partir de principes fondamentaux ou primaires. Par contre, il n'est pas rare qu'elle essaie de se remémorer une situation similaire à laquelle elle fait face. En fonction des souvenirs qu'elle réussit à extraire de sa mémoire, elle essaiera d'adapter la stratégie autrefois utilisée à la nouvelle situation. Cette approche à base de cas peut être une avenue prometteuse à la génération d'explications. De la même façon qu'une personne utilise son expérience dans le but de résoudre un problème, cette expérience pourrait devenir le point de départ pour une nouvelle explication.

Cependant, se rappeler de vieilles explications peut parfois être très facile, mais également extrêmement difficile à d'autres moments. Lorsqu'une personne se souvient d'un épisode qui semble *a priori* inapplicable, il faut un effort de créativité remarquable à une seconde personne pour suggérer une explication valable à partir de cet épisode. Pour qu'un programme puisse faire face à la réalité souvent changeante et complexe, il doit faire preuve d'apprentissage et de créativité. Ces systèmes doivent être capables de réviser et modifier une vieille explication pour expliquer un événement qui n'est pas similaire aux situations rencontrées dans le passé. Afin de profiter de ces nouvelles explications générées, le système doit également être en mesure de se rappeler des explications produites. Il sera donc nécessaire d'indexer les explications de façon à ce qu'elles soient toutes accessibles le plus rapidement possible. Dans le but d'expliquer cette approche plus en détails, nous nous baserons sur le système SWALE [Schank & Leake, 89].

4.5.1 LE PROJET SWALE

SWALE est un système qui produit de nouvelles explications par l'extraction d'anciennes explications (d'une base propre à un système) pour ensuite les modifier afin qu'elles puissent s'appliquer à la nouvelle situation. Le but du système est de trouver une explication qui pourrait aider l'utilisateur à comprendre un événement anormal. Le système possède une base de cas représentant des histoires à propos d'animaux ou d'humains décédés dans d'étranges circonstances. Par exemple, si l'on donne au système un anormal tel que la mort d'un cheval de course en pleine santé, le système cherchera des explications dans une base comportant d'autres événements anormaux. La représentation des explications en mémoire doit être bien déterminée. Les structures de connaissances doivent être assez flexibles pour supporter un processus d'adaptation. Cette structure a pour nom XP (eXplanation Patterns). À partir de cette structure, il est possible d'appliquer directement des sections d'une vieille explication générée. Par le même procédé, la structure doit permettre d'identifier et de remplacer les sections de l'explication qui ne sont pas applicables. De cette façon, la réutilisation de la banque d'explications présente en mémoire est maximisée. De façon à ce que le système puisse être en mesure de reconnaître lorsqu'une XP s'applique à une anomalie précise, chaque XP doit également avoir une représentation des relations qui existent à l'intérieur d'une base d'explications afin d'aller chercher des précisions indispensables aux explications.

Nous présenterons maintenant l'algorithme général derrière le système SWALE afin de mieux saisir l'utilité de cette technique.

1. **Détection d'anomalie**

   Essayer de trouver une histoire correspondante dans la base d'histoires. Lorsque SWALE fait la lecture d'une histoire, *il essaie d'associer chaque fait (entrée) à* une connaissance déjà existante. Si un fait ne peut être associé, il devient une anomalie.

   Si la recherche est un succès, TERMINÉ; sinon une anomalie a été détectée.

2. **Recherche d'une XP**

   Recherche d'une XP qui pourrait être appliquée à l'anomalie afin de l'expliquer.

3. **Acceptation de l'XP**

   Essayer d'appliquer l'XP.
   Si l'XP s'applique, alors aller à l'étape 5.

4. **Modification de l'XP**

   S'il est impossible d'appliquer l'XP à l'anomalie afin de la comprendre, essayer de trouver une nouvelle XP qui pourrait s'appliquer.
   Si la recherche d'une nouvelle XP est un succès, alors retourner à l'étape 3.

5. **Généralisation**

   Si réussite, intégrer les résultats dans la base de connaissances en vue du futur.

4.6 MODÈLE EES (EXPLANABLE EXPERT SYSTEM)

Un problème souvent rencontré lors de la génération d'explications, est le manque d'information sur la façon dont le système expert a été construit. Lorsque vient le temps d'expliquer le raisonnement effectué, il est fort probable qu'il sera très difficile de justifier les actions accomplies. Le modèle EES [Swartout & Moore, 93] a été conçu afin de pouvoir capter l'information sur la conception du système expert et ainsi être capable de générer des explications adéquates.

La conception d'un tel système débute par la construction d'une base de connaissances de haut niveau. En collaboration avec des experts du domaine visé, la conception de cette base de connaissances doit comporter les faits du domaine, la terminologie du domaine ainsi que les stratégies générales à la résolution de problèmes. Cette base de connaissances peut être vue comme une spécification abstraite du système expert qui sera construit. La figure 6 présente le modèle du système EES.

Par la suite, il y a construction de façon automatique du système expert depuis la spécification de haut niveau. Cette construction est effectuée à l'intérieur même du système par un module spécifique à la construction du système expert. Pendant que ce module construit le système expert, il enregistre à la fois toutes les décisions de conception à l'intérieur d'un historique. Le code produit par ce module de construction automatique est lié directement à l'historique de la conception. À l'exécution, le code du système expert est exécuté par un interpréteur. Lorsque l'utilisateur demande des explications sur le comportement du système expert, le module ayant comme responsabilité la génération d'explications entre en jeu. Ce dernier utilise l'historique de conception pour déterminer quelles sont les décisions de conception ayant mené à l'implantation du système. Il produit ainsi des explications qui n'indiquent pas seulement à l'utilisateur ce qui arrive, mais qui justifient également ces actions. Ces justifications, qui portent sur les actions prises par le système, s'accordent en fonction des décisions de conception et des connaissances comprises dans la base de haut niveau. Ce modèle capture donc le raisonnement qui sous-tend les actions posées par le système.

Une particularité de ce système est qu'il dérive ses actions directement de la spécification abstraite concernant les méthodes générales et les propriétés du domaine. Puisque ces liens sont automatiquement dérivés, le système demeure fidèle à lui-même tout au long du processus. Nous présenterons maintenant plus en détails la base de connaissances et le module de construction automatique de ce modèle.

4.6.1 LA BASE DE CONNAISSANCES DU MODÈLE EES

Cette base fait la distinction entre trois types de connaissances différentes :

- Les connaissances reliées à la terminologie;
- Les connaissances reliées au domaine d'expertise;
- Les connaissances reliées à la résolution de problème.

Les connaissances reliées à la terminologie définissent les termes d'un domaine particulier et donnent à ces termes une sémantique particulière. Ces termes sont utilisés dans la construction de plans pour la résolution de problèmes et dans le modèle du domaine. Ce dernier modèle contient les éléments du domaine qui spécifient comment les différents objets de ce domaine sont reliés entre eux. Cependant, ce modèle ne précise pas comment résoudre les problèmes liés au domaine. Ceci est réalisé par les connaissances reliées à la résolution de problèmes. Ces connaissances sont représentées par des plans ou stratégies visant un but particulier. Chaque stratégie inclut une description de ces possibilités et une séquence abstraite des actions posées afin d'atteindre le but visé.

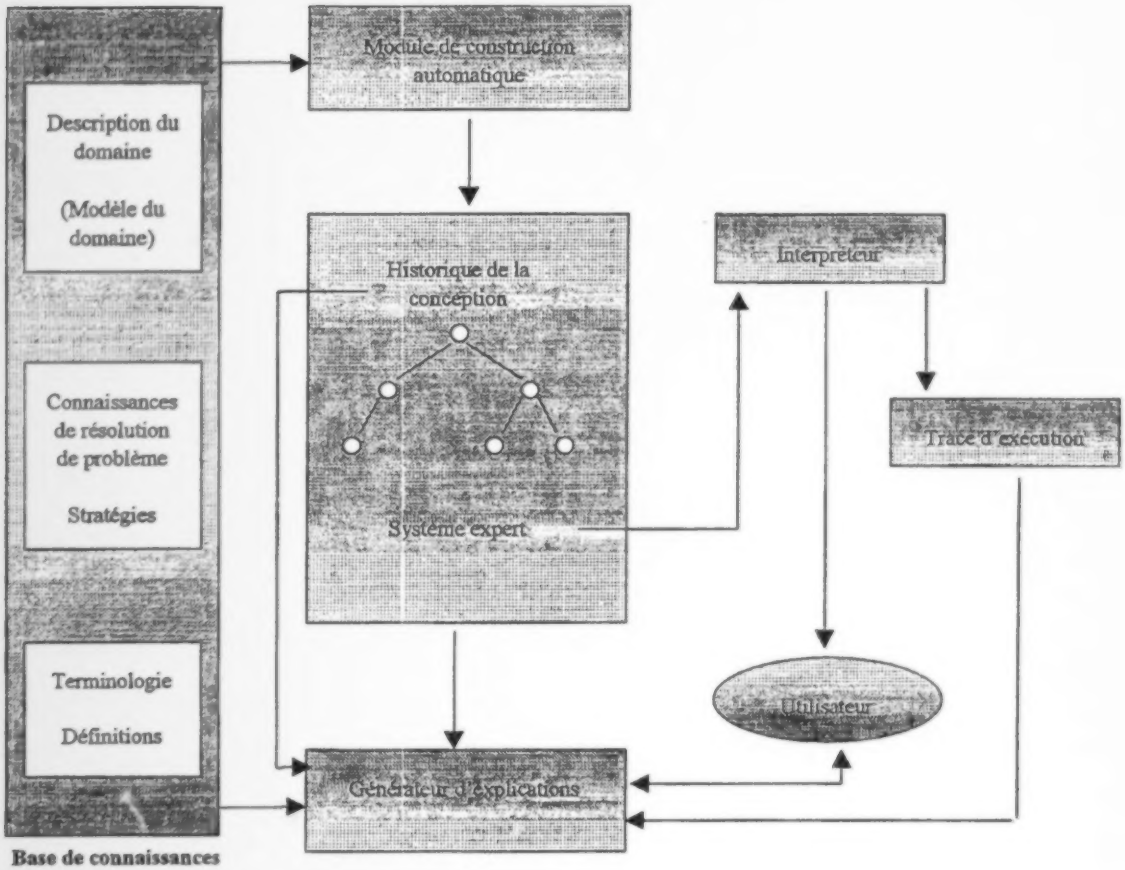

Figure 6 Modèle du système EES [Swartout & Moore, 93]

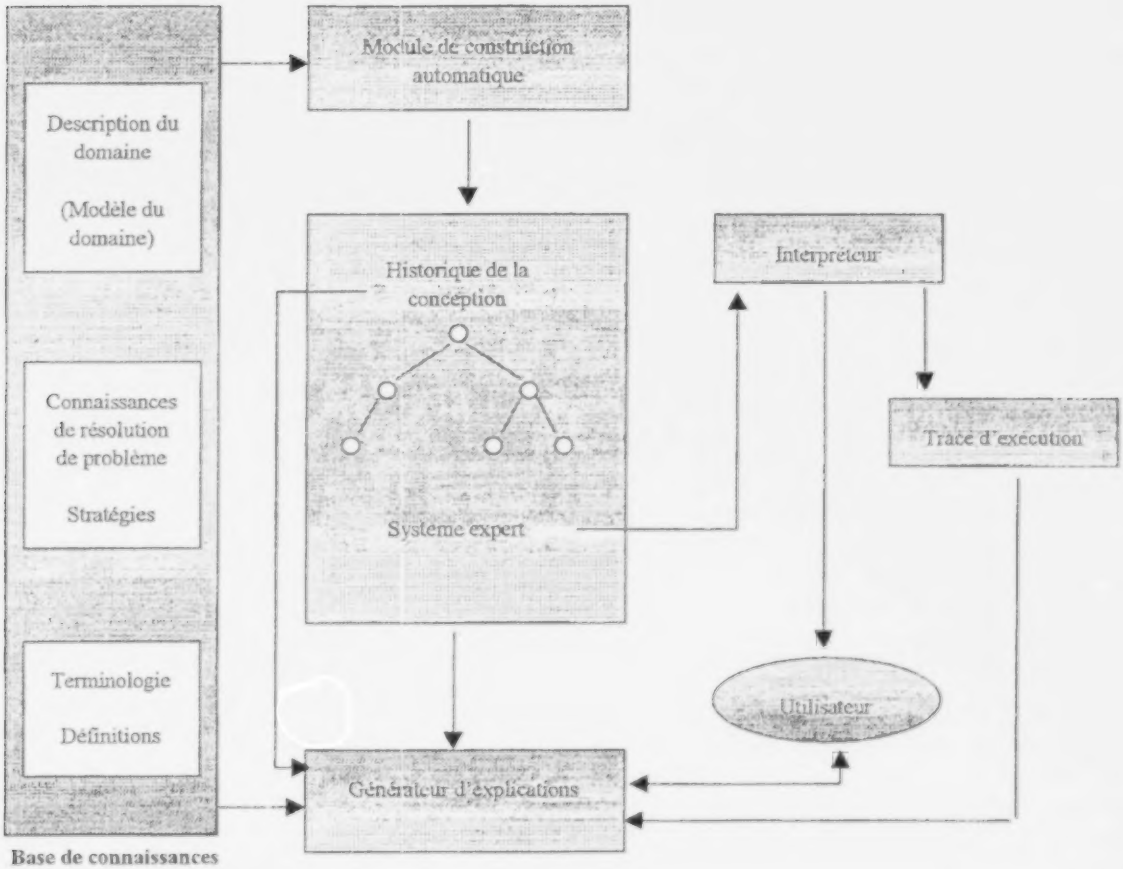

Figure 6 Modèle du système EES [Swartout & Moore, 93]

4.6.2 MODULE DE CONSTRUCTION AUTOMATIQUE

Le module de construction automatique fonctionne sur le principe de raffinements successifs. Initialement, le module débute avec un but général qui indique ce que le système expert est supposé accomplir. Le module fait alors une recherche à travers les connaissances reliées à la résolution de problème pour essayer de trouver une stratégie qui possède les critères nécessaires à l'obtention du but visé. Si aucune stratégie n'est trouvée, le module tente de reformuler le but en un nouveau but ou ensemble de buts qui ont la même intention que le but original, mais est exprimée de façon différente. Ce processus de reformulation est important vis-à-vis les explications, car cette reformulation engendre le raisonnement nécessaire à la justification des actions accomplies par le système.

4.7 MODÈLE BASÉ SUR LA RECONSTRUCTION DES EXPLICATIONS

Lorsque l'on demande à l'être humain d'expliquer un raisonnement complexe, il est improbable qu'il l'explique exclusivement en termes du processus de résolution utilisé. Il tentera plutôt de construire une histoire qui appuie le problème résolu. Cette histoire reflétera une suite d'explications, qui ne sera pas nécessairement identique à la suite du raisonnement associé [Wick & Thompson, 92]. La suite de raisonnement est la trace du système expert lors de la résolution du problème. Dans la construction d'une suite d'explications, il est possible d'ajouter de l'information afin de créer un support additionnel. De plus, certaines évidences non utilisées lors de la résolution du problème sont également introduites dans les explications. Ceci inclut non seulement les causalités sous-jacentes, mais aussi de nouveaux symptômes aidant à supporter la conclusion finale. La caractéristique première de ce modèle est la rupture du lien entre le moteur servant à la génération des explications et celui exécutant le processus de résolution de problèmes. Par la rupture de ce lien, le système d'explication a la liberté de créer plus naturellement un support solide aux actions et conclusions du système expert.

Par l'entremise de cette méthode, des méthodes et une base de connaissances propre à la résolution de problèmes sont utilisées au début pour résoudre le problème. Une fois le problème résolu, une seconde base de connaissances, celle-ci orientée vers l'explication, assemble une histoire plausible qui justifie la conclusion atteinte. Cette histoire peut utiliser toutes les composantes de la trace de résolution ou simplement quelques-unes d'entre elles. La seconde base de connaissances n'a pas à être orientée vers la résolution d'un problème comme la première. Après avoir reçu la conclusion du système expert, elle doit construire une suite d'explications à partir des entrées jusqu'à la conclusion, tout en considérant la trace de résolution. Il est donc possible de tailler la première base de façon plus efficace envers la résolution, sans prendre en considération les explications qui devront être générées par la suite [Wick & Thompson, 92].

4.7.1 LE SYSTÈME REX

Le système REX (Reconstruction EXplainer) [Wick & Thompson, 92] est le prototype d'un système d'explication entièrement fonctionnel basé sur l'approche de la reconstruction des explications. REX est conçu afin de fournir des explications sur comment un système expert évolue et comment des données d'entrées d'un problème permettent de générer la conclusion atteinte par ce dernier. Le principal atout du système REX est qu'il permet de valider le degré de couplage entre le chemin que le système expert emprunte pour en arriver à une conclusion et les explications qui y sont reliées. Si quelques paramètres sont modifiés, le système REX peut être obligé à suivre le raisonnement du moteur d'inférences de très près. Par contre, il possède également la liberté de générer des explications qui justifient la conclusion, mais ne présente pas tout le raisonnement du système. Cette validation permet de choisir le type d'explication que l'on veut générer selon l'utilisateur du système expert. En effet, un critère qui influence grandement la nature d'une explication est le type d'audience. Nous devons tenir compte de la personne à qui s'adresse l'explication afin d'en définir le contenu et l'organisation. Trois principales audiences peuvent être ciblées :

- Ingénieur des connaissances;
- Expert du domaine;
- Utilisateur.

L'ingénieur des connaissances sera intéressé aux opérations internes du système expert. Un expert du domaine utilisera généralement une explication afin de déterminer ce que le système connaît ainsi que le niveau des connaissances du système. Finalement, l'utilisateur du système est intéressé à une meilleure compréhension des recommandations données par le système.

Selon l'approche du système REX, un système expert conventionnel est d'abord utilisé pour solutionner un problème. Le système garde la trace du processus de résolution qui comporte plusieurs éléments tels que [Swartout & Moore, 93] :

- Les différentes étapes temporelles du processus;
- Les données utilisées;
- Les hypothèses générées qu'elles soient acceptées ou refusées.

Cette trace est ensuite passée à travers un filtre qui, basé sur des contraintes liées au type d'audience, détermine l'information contenue par la trace du raisonnement disponible au système qui génère les explications. Quatre options composent les contraintes du problème:

- Aucune composante : Aucun élément de la trace venant du processus passe à travers le filtre, exception faite de la conclusion.

- Composantes directes : Seuls les faits qui composent la trace du processus passent le filtre. Toute hypothèse générée par le système est éliminée.
- Composantes indirectes : Seules les hypothèses générées par le système passent à travers le filtre.
- Composantes directes et indirectes : La trace complète du processus passe à travers le filtre.

Ce filtre est l'un des mécanismes du système REX qui permettent de contrôler le degré de couplage entre le processus d'inférence et les explications. Plus il y a d'informations qui passent par le filtre, plus la suite d'explications est proche de la suite du raisonnement lors de la résolution. La figure 7 montre une généralisation du système basée sur la reconstruction des explications.

L'information qui survit au processus de filtrage est jumelée à une base appelée spécification des connaissances. Cette base définit une interface entre les connaissances utilisées par le système expert et celles utilisées par le système d'explication. Cette spécification est utilisée pour représenter une base commune entre la base de connaissances du système expert et celle du système d'explication. L'idée est d'utiliser une spécification de haut niveau du domaine de telle sorte que les détails reliés au raisonnement du système peuvent être ignorés et l'essence de ce que le système a réalisé comme raisonnement est sauvegardée. De plus, cette spécification permet d'effectuer un second filtrage au niveau des détails reliés au raisonnement du système expert. Ensuite, cette spécification passe par le générateur d'explications sous certaines contraintes sur la solution. Ces contraintes sont du même type que celles appliquées au problème, elles influencent le degré d'utilisation des traces de raisonnement présentes dans la spécification liée au processus de résolution. Elles ne font que renforcer ou alléger les contraintes du problème.

Le générateur d'explications a la responsabilité de chercher à travers la base d'explications les relations et méthodes qui soutiennent l'information sélectionnée dans la spécification des connaissances. Le résultat est une suite d'explications qui sera éventuellement présentée à l'utilisateur sous une forme plus compréhensible. Les contraintes sur la solution agissent comme contrôle sur le niveau de liberté donnée au système d'explication. La base d'explications est composée de connaissances orientées vers l'explication. Cette base est le point central du processus d'explications.

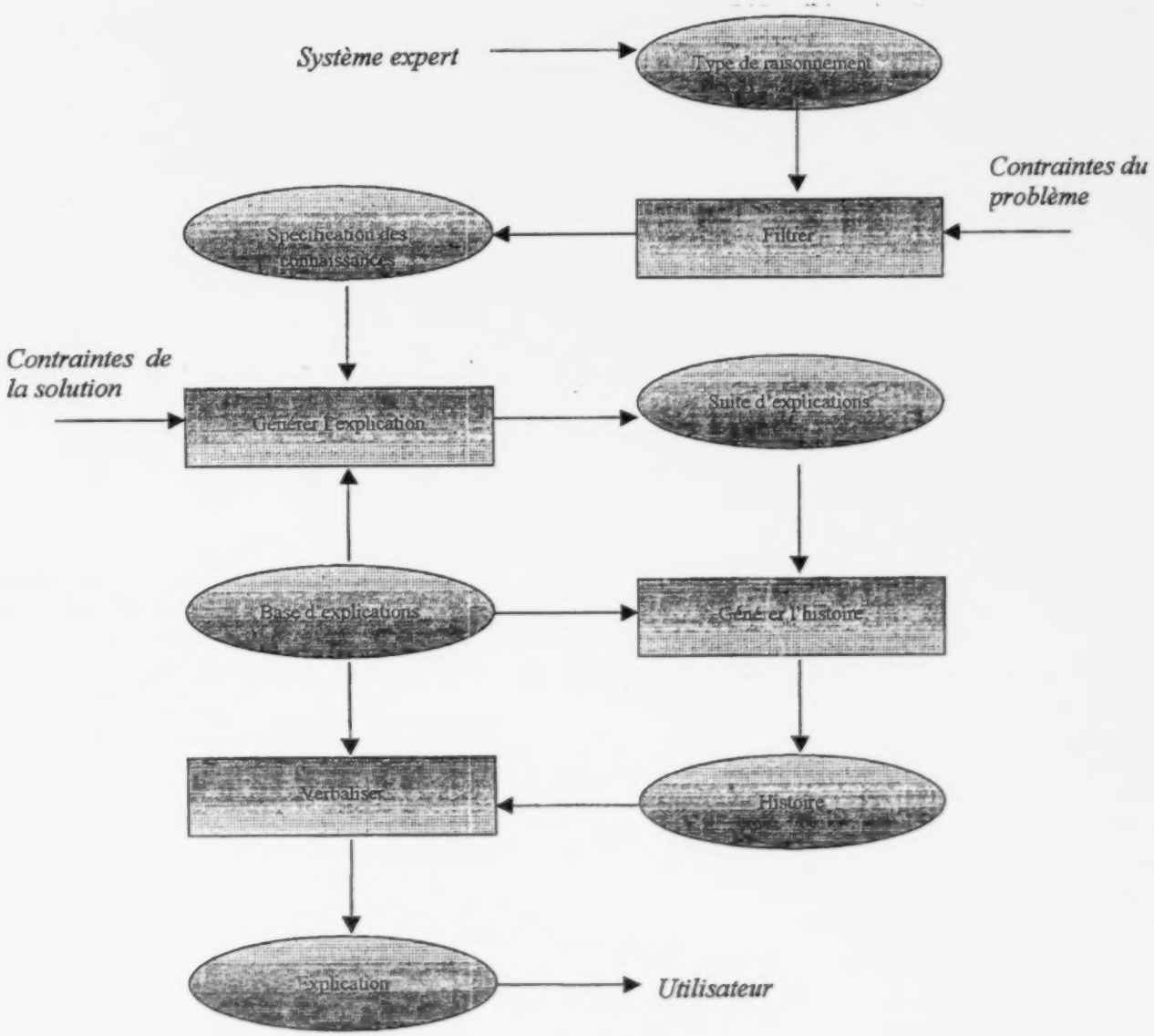

Figure 7 Modèle du système d'explication REX [Wick, 93]

Des recherches sur la résolution de problèmes chez l'humain ont démontré que le type de connaissances utilisées à la résolution d'un problème et le type de connaissances utilisées à la description de la solution sont différents. Premièrement, la base de connaissances orientée vers l'explication possède une structure beaucoup plus générale que celle

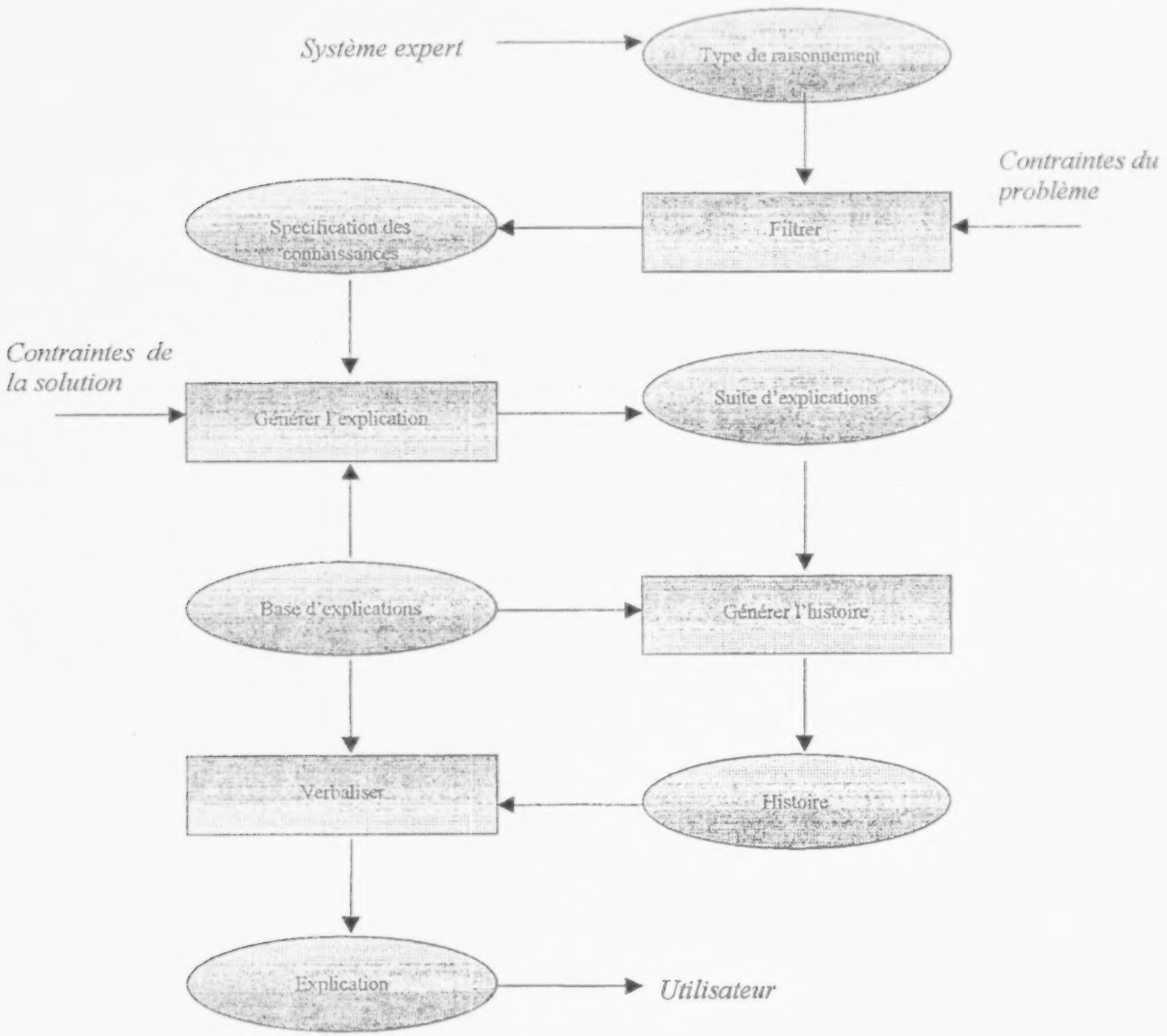

Figure 7 Modèle du système d'explication REX [Wick, 93]

Des recherches sur la résolution de problèmes chez l'humain ont démontré que le type de connaissances utilisées à la résolution d'un problème et le type de connaissances utilisées à la description de la solution sont différents. Premièrement, la base de connaissances orientée vers l'explication possède une structure beaucoup plus générale que celle

orientée vers la résolution d'un problème lié au système expert. En particulier, la base associée aux explications n'a pas de *ligne de pensée* qui est généralement une restriction chez les systèmes experts [Wick, 93].

Une fois que le générateur d'explications a construit une suite d'explications, celle-ci est dirigée vers le générateur d'histoire. Ce dernier est responsable de prendre comme entrée la suite d'explications et de l'organiser dans un flot cohérent depuis les données d'entrées jusqu'à la conclusion. Cette suite d'explications est présentée comme une histoire qui raconte comment le système expert a parcouru une série d'hypothèses de l'état initial à la conclusion finale. L'histoire résultante est ensuite dirigée vers l'étape de verbalisation qui traduit l'histoire vers les explications finales qui seront données à l'utilisateur.

4.8 LE RAISONNEMENT PROFOND

Le raisonnement explicatif [Kassel, 89] est un raisonnement à part entière, concerné notamment par les questions suivantes :

- Que doit-on expliquer : la solution trouvée, le raisonnement dans son ensemble ou bien les étapes importantes?
- Quelles sont les attentes de l'interlocuteur : désire-t-il savoir comment une méthode a été appliquée ou connaître les raisons du choix de cette méthode?
- Comment expliquer : doit-on commenter, justifier ou simplement donner une idée générale?

Ces questions mettent en évidence un aspect particulier du raisonnement explicatif que chacun d'entre nous peut constater de façon intuitive, à savoir qu'il peut être plus ou moins complexe, et donc plus ou moins profond, selon l'objectif et la qualité des explications que nous souhaitons obtenir.

Le raisonnement explicatif peut se décrire assez naturellement en termes de raisonnement de surface et de raisonnement profond. Un problème explicatif simple, par exemple présenter une étape d'un raisonnement ou sa conclusion, peut se résoudre simplement sans qu'il soit nécessaire de comprendre le raisonnement. Des études ont démontré que le processus psychologique du raisonnement explicatif montre bien d'ailleurs que pour de tels problèmes simples, nous [Kassel, 89] disposons de *patterns* d'explication prêts à l'emploi. Un problème explicatif plus complexe, à savoir justifier la conclusion ou certains points du raisonnement, présenter une synthèse du raisonnement, ou montrer que plusieurs raisonnements étaient possibles pour arriver à la même conclusion, nécessitera au contraire un raisonnement profond exigeant une compréhension en profondeur du raisonnement à expliquer. Il est intéressant d'observer que la complexité du problème peut aussi être liée à une exigence de qualité, comme le fait de fournir des explications pédagogiques tenant compte du niveau de connaissances de l'interlocuteur.

4.8.1 L'ARCHITECTURE DU SYSTÈME CQFE

Le système explicatif CQFE (Ce Qu'il Fallait Expliquer) [Kassel, 89], développé à l'Université d'Orsay, a été validé sur une application de diagnostic technique et illustre parfaitement l'intérêt d'une approche de raisonnement profond pour construire un module d'explication.

L'objectif de ce module est de présenter à l'utilisateur, en fin de session, une synthèse du raisonnement suivi par le système expert. Pour cela, le module d'explications opère un nouveau raisonnement comportant trois étapes distinctes : le raisonnement du système est tout d'abord structuré, puis il est simplifié et en dernier lieu présenté à l'utilisateur. La première étape consiste simplement à exploiter, de façon dynamique, un environnement objets pour mettre en évidence les objets du raisonnement.

Pour fournir une vue globale du raisonnement, il est hors de question de présenter toutes les inférences effectuées. Au cours de la seconde étape, le module d'explications dégage donc les inférences lui paraissant les plus importantes, les autres étant jugées non pertinentes pour l'explication. Pour cela, il utilise deux types de connaissances profondes [Gréboval, 96] :

- Une théorie du domaine, appelée modèle conceptuel, contient une description des objets et concepts qui structurent le savoir-faire (règles) de l'expert. Ce modèle contient notamment la topologie du matériel à diagnostiquer ainsi que les propriétés générales des différents éléments. Par exemple, le fait que tel moteur soit de type électrique et qu'il possède, comme tout appareil électrique, une alimentation électrique et des fusibles. D'autres connaissances sont représentées dans le modèle conceptuel, notamment un modèle fonctionnel du matériel; seules les connaissances que nous venons d'indiquer sont utilisées lors de ce raisonnement.

- Un savoir-faire explicatif, composé des connaissances C1 et C2, spécifie comment simplifier le raisonnement :
  - C1 : une inférence dont la conclusion est déductible à l'intérieur du modèle conceptuel est non pertinente.
  - C2 : supprimer du raisonnement les inférences jugées non pertinentes.

L'explication fournie par CQFE s'adresse à un technicien. Le module d'explications fait ici l'hypothèse que ce technicien est familier avec les connaissances contenues dans le modèle conceptuel et qu'il est au contraire intéressé par le savoir-faire en diagnostic du SE, ce qui justifie C1. Ainsi, les inférences concluant sur un fait supposé déjà connu de l'interlocuteur sont supprimées (Connaissance C2).

Au cours de la troisième étape du raisonnement explicatif, la présentation du raisonnement structuré et simplifié, un nouveau raisonnement profond est mis en œuvre qui utilise cette fois un modèle fonctionnel du matériel.

Ce raisonnement consiste à distinguer parmi les composantes sur lesquelles le système expert a raisonné les composants défectueux de ceux ayant normalement fonctionné (première et seconde parties reproduites en caractères gras dans l'encadré suivant afin de reconstituer le scénario de la panne et de présenter le diagnostic).

Résume → (arrêt manutention)
Les faits suivis d'un astérisque sont ceux que vous avez affirmés:
1→ axe-poulie-gauche
rupture: vrai (*)
Ce composant est à l'origine de la panne
Justification du fait (arrêt manutention)
2→palonnier-d-équilibrage
détecte-déséquilibre: vrai
3→frein-de-service
serre: vrai
1 entraîne 2; 2 entraîne 3
On en déduit donc le fait

Exemple du dialogue explicatif avec CQFE

Nous retrouvons dans la description du raisonnement explicatif de CQFE la connaissance et un savoir-faire explicatif. La « théorie explicative » contient la théorie du domaine (les modèles du domaine) utilisée par le raisonnement du système expert. Ceci est logique dans la mesure où une bonne explication nécessite une compréhension profonde du raisonnement. Elle contient, en outre, d'autres informations [Kassel, 89] :

- Un historique du raisonnement : l'origine de déduction de chaque fait doit être mémorisée;
- Un modèle de raisonnement du système expert : ce modèle est nécessaire pour reconstituer et analyser la trace du raisonnement;
- Un modèle de l'interlocuteur, qui peut avoir plusieurs dimensions : désirs, attentes, niveau de connaissances de l'interlocuteur.

Le savoir-faire explicatif est une expertise spécifique comprenant notamment :

- Des stratégies de rhétorique; par exemple, attendre de savoir si l'interlocuteur a compris avant de fournir une justification;
- Des stratégies de pédagogie; par exemple, dans un contexte d'enseignement, si l'élève demande à connaître la méthode utilisée et que l'on désire qu'il continue à chercher, lui fournir simplement quelques indications.

La qualité de l'explication d'un raisonnement dépend à la fois du modèle de ce raisonnement et de la nature du raisonnement explicatif.

4.9 DES EXPLICATIONS À L'AIDE D'AGENTS INTELLIGENTS

Cette approche est centralisée sur l'utilisation d'agents intelligents qui, par leur propre autonomie, ont la capacité d'expliquer et de justifier leurs actions. Ceci est particulièrement vrai et utile lors de la simulation de situations précises telles qu'une attaque navale ou aérienne. L'intérêt pour de telles simulations est grandissant envers l'évaluation de missions, de tactiques et même de doctrines [Johnson, 94].

La validation ou l'évaluation d'un comportement est difficile à effectuer lorsque la position de l'observateur est interne. Un système composé de plusieurs agents (ou système multi-agents), est un meilleur moyen pour évaluer ses propres actions posées. Il est en mesure de déterminer les motivations d'une décision en reprenant la situation établie lors de la décision originale et d'y injecter les mêmes variantes initiales. À travers cette expérimentation, les agents sont en mesure de découvrir quels sont les facteurs qui ont mené à la décision de même que les options qui auraient pu être envisagées lors d'une situation quelque peu différente. Ces agents apprennent également à reconnaître les situations similaires où les mêmes décisions peuvent être prises pour les mêmes raisons.

Une seconde approche est la vue de l'agent par rapport à l'utilisateur qui manipule le système d'aide à la décision. Une problématique majeure est que les explications sont souvent dédiées uniquement à un expert du domaine d'application. Un modèle intéressant serait un système multi-agents où chaque agent possède une certaine abstraction des connaissances d'explication. Ainsi, à l'aide d'une coopération [Tadié *et al*, 96] entre les différents agents intelligents du système, plusieurs types d'explications, tous transportant la même information, pourraient être générés. Donc, si un agent représente le profil de l'utilisateur, une dynamique pourrait être formée à l'intérieur du système dans le but de générer des explications conçues idéalement pour l'utilisateur.

D'une manière générale, la contribution à l'aide d'agents permet de distribuer à plusieurs niveaux à la fois le type d'explication, le niveau de détails et par conséquent la granularité des connaissances impliquées, enfin les objectifs vis-à-vis d'un type d'utilisateur. Ces approches font actuellement l'objet de recherches au laboratoire Héron de l'Université de Montréal [Tadié *et al*, 98], [Frasson, 98].

5. ANALYSE DES AVANTAGES ET COÛTS DES TECHNIQUES D'EXPLICATION

Ce chapitre a pour objectif d'identifier les avantages et les coûts reliés à l'intégration de mécanismes d'explication automatique dans les systèmes d'aide à la décision. Cette analyse est générique et n'est donc pas spécifique à une application ou à une technique d'explication particulière. C'est pourquoi les avantages et les coûts présentés dans ce chapitre doivent être mis en contexte avec chaque application pour en évaluer finement leur importance réelle. La section 5.3 présente un tableau synthèse des principaux avantages et coûts relatifs à l'ajout de techniques d'explication.

5.1 ANALYSE GÉNÉRALE DES AVANTAGES DES TECHNIQUES D'EXPLICATION

L'analyse des avantages relatifs à l'implantation de techniques d'explication dans un système d'aide à la décision a été regroupée en cinq thèmes spécifiques de manière à simplifier leur présentation. Ces thèmes sont les suivants :

- Qualité des recommandations;
- Aspects psychologiques;
- Processus de prise de décisions;
- Analyse du comportement du système;
- Gestion des connaissances.

Cette section présente une description des avantages identifiés dans la littérature et lors des expériences accumulées dans les différents projets réalisés chez Novasys.

5.1.1 QUALITÉ DES RECOMMANDATIONS

Ce groupe d'éléments vise à présenter les principaux avantages de l'ajout d'explication sur la qualité des recommandations fournies par un système d'aide à la décision.

5.1.1.1 Meilleure compréhension des problèmes complexes

Les utilisateurs de systèmes d'aide à la décision sont souvent confrontés à des problèmes comportant un volume important de connaissances et de relations entre les connaissances. Ces cas de problèmes complexes ne sont pas simples à comprendre rapidement et il est difficile de s'assurer qu'aucun élément important n'ait pas été oublié dans le raisonnement. Selon [Wick et Thompson, 92], les explications fournies par un système d'aide à la décision facilitent la compréhension de ces problèmes complexes.

5.1.1.2 Augmentation du niveau de compréhension des recommandations données à l'utilisateur

Une grande majorité des systèmes d'aide à la décision développés dans le passé se limitaient à fournir une recommandation à l'utilisateur afin qu'il puisse agir en conséquence. Toutefois, les recommandations proposées par ces systèmes ne correspondaient pas toujours à quelque chose qui avait du sens pour l'utilisateur.

L'ajout d'explications jumelées avec ces recommandations a pour avantage de donner les raisons, les motifs et la nature du raisonnement et des comportements du système qui l'ont amené à proposer ces recommandations. Les explications peuvent aussi préciser le contexte du problème, les relations entre les éléments, la chronologie des comportements, les solutions rejetées et les hypothèses vérifiées par le système. Tous ces éléments favorisent une meilleure compréhension des recommandations.

5.1.1.3 Diminution des ambiguïtés

L'ajout d'explications permet de préciser et de compléter les recommandations de manière à éliminer les possibilités d'interprétations mauvaises, incertaines, contradictoires ou de nature qui diffère de la réalité. Elles peuvent faire en sorte que les recommandations ne laissent aucun doute sur la décision à prendre.

5.1.1.4 Explication adaptable selon les besoins et contraintes des utilisateurs

Les recommandations d'un système d'aide à la décision sont souvent très concises et présentent une solution sans préciser comment le système a fait pour y arriver. Les explications sont alors un ajout important pour détailler le raisonnement et faire en sorte que l'utilisateur comprenne bien cette solution.

Toutefois, ces explications ne s'adressent pas nécessairement à une même personne d'où l'avantage important de les personnaliser afin qu'elles puissent tenir compte des connaissances de l'utilisateur [Wick et Thompson, 92], [Swartout et Moore, 93], de ces préférences [Goguen et al., 83], de ces prédispositions face au problème [Goguen et al., 83] ou du niveau de dialogue nécessaire. De plus, les explications peuvent aussi être adaptées en fonction de besoins ponctuels durant le raisonnement (en mode dialogue [Moore et Swartout, 89]) ou le besoin d'une explication graduelle ou avec exemples. Enfin, plusieurs stratégies explicatives sont susceptibles d'être nécessaires selon la complexité du problème comme, par exemple, l'ajout d'éléments explicatifs multimédias (image, son, animation, etc.).

5.1.2 ASPECTS PSYCHOLOGIQUES

Le processus de prise de décision est souvent un processus complexe et les impacts d'une mauvaise décision sont parfois très importants et psychologiquement exigeants pour le décideur. Dans un tel contexte, les informations qui lui sont fournies sont capitales pour la réalisation de ses fonctions. Les éléments examinés dans cette sous-section proposent différents avantages psychologiques relatifs à l'ajout d'explication dans un système d'aide à la décision.

5.1.2.1 Augmentation du niveau de confiance

Les systèmes d'aide à la décision sont développés de manière à intégrer un ensemble de connaissances (savoir et de savoir-faire) qui ne sont pas nécessairement accessibles par l'utilisateur du système. Ainsi, le fait de compléter les recommandations par des explications ou d'offrir la possibilité d'expliquer chacune des actions du système durant son raisonnement peut aider l'utilisateur à mieux prendre conscience des capacités et du niveau d'expertise du système et de la qualité de ces recommandations.

Du point de vue psychologique, la confiance concerne les situations dans lesquelles nous nous en remettons à autrui pour l'obtention d'un résultat. En tant qu'attitude, la confiance peut alors être définie comme un ensemble de croyances en certaines qualités (connaissances et expérience) ou intentions de l'autre (le système dans ce cas), ([Butler, 91]; [Mayer et al., 95] et [Mishra, 96]; [Griffin, 67]; [Jones, James et Bruni, 75]; [McAllister, 95]) in [Couteret, 98].

Dans ce sens, les explications dans les systèmes d'aide à la décision sont importantes parce qu'elles amènent les utilisateurs à avoir plus confiance au système [Waterman, 86] ; [Haton et al., 91] et les aident à comprendre les conséquences de ses conclusions [Wick, 93]. Elles augmentent la vraisemblance et le crédit (mérite, valeur) à donner à la qualité des recommandations [Swartout et Moore, 1993].

Cette crédibilité est importante pour la prise de décision de l'utilisateur du système car selon [Hovland et Weiss, 51] in [Ghiglione et al., 92], une même information peut influencer différemment les sujets selon qu'elle est attribuée à une source très crédible (forte influence) ou peu crédible (influence faible ou nulle). Donc, sans explications les décideurs pourraient douter de la crédibilité d'une recommandation et la refuser. Ainsi, plus le système justifie et détaille son raisonnement, plus l'utilisateur est enclin à avoir confiance aux recommandations lors de sa prise de décision et encore plus dans des situations de grand stress ou lorsque les impacts de ces décisions sont très importants. Elles servent alors à persuader l'utilisateur à agir dans le même sens que le système [Waterman, 86] ; [Swartout et Moore, 93].

5.1.2.2 Meilleure perception du comportement d'aide et de coopération fournie par le système

L'ajout d'explication dans un système d'aide à la décision a pour objectif, entre autres, d'assister et de soutenir l'utilisateur dans son processus de prise de décision. Les explications donnent les moyens pour s'assurer que les recommandations sont bien comprises et que le raisonnement est de qualité. Le dialogue explicatif est un exemple significatif démontrant un comportement d'aide, de support et de coopération du système car l'utilisateur peut, durant le raisonnement, demander des précisions pour l'aider à comprendre ou pour confirmer l'efficacité du comportement du système.

5.1.2.3 Possibilité de diminution de la perception de contrôle par le système

Les explications fournies par le système peuvent donner l'impression qu'il donne toute l'information qu'il possède, qu'il n'a rien à cacher et qu'il donne plus de moyens pour le critiquer et valider son efficacité. Dans ce sens, des fonctions d'explication rendent un système expert plus transparent [Swartout et Moore, 93]. Ainsi, les utilisateurs ayant toute l'information disponible ont plus l'impression d'avoir le contrôle sur la prise de décision que lorsque aucune explication n'a été fournie par le système.

5.1.3 PROCESSUS DE PRISE DE DÉCISION

Dans un système d'aide à la décision, la décision finale est le but ultime. Cette décision est souvent soumise à de nombreuses contraintes et elle se doit d'être de qualité. Les techniques d'explication implantées dans de tels systèmes peuvent être très bénéfiques pour la qualité de ces décisions. Voici une description des principaux avantages de ces techniques pour l'efficacité du processus de décision.

5.1.3.1 Possibilité de rétroaction suite aux explications

La possibilité d'obtenir en cours de raisonnement (en mode dialogue) des explications du système d'aide à la décision offre à l'utilisateur les avantages suivants : demander plus d'explications, éclaircir certains éléments, ajouter ou modifier des éléments d'information selon les besoins. Ainsi, les informations données au système ont plus de chance d'être exactes et de refléter la réalité du problème et des faits. Sans les explications, il pourrait s'avérer difficile de valider la justesse et la pertinence des réponses données par l'utilisateur et la qualité du raisonnement du système.

5.1.3.2 Réduction du temps alloué à la prise de décision

Associer des explications au raisonnement d'un système d'aide à la décision permet de mieux comprendre les recommandations, mieux évaluer leur qualité et ainsi pouvoir agir en confiance plus rapidement. Dans ce sens, le processus de prise de décision peut tirer avantage de l'ajout d'explication.

5.1.3.3 Augmentation de la qualité des décisions

Sans explication, les décideurs ont moins de moyens pour s'assurer de la qualité des informations saisies et des résultats d'un système. Toutefois, ils doivent agir en fonction de ces recommandations ou choisir une toute autre voie. Si les informations disponibles ne permettent pas de déceler un problème au niveau des informations fournies ou du comportement du système, des erreurs peuvent alors survenir dans les décisions qui suivent. Ainsi, la qualité des décisions des utilisateurs pourrait être affectée par la présence ou non d'explication.

De plus, le fait d'avoir des explications permet aux utilisateurs de prendre des décisions plus éclairées, cohérentes et uniformes dans un même contexte. En effet, lorsque le système demande des informations complémentaires à l'utilisateur, celui-ci pourrait ne pas interpréter cette question de la même façon qu'un autre utilisateur et fournir une réponse différente qui aurait pour impact de conclure différemment à un même problème. Dans ce cas, les explications peuvent aider à augmenter la qualité des décisions en diminuant les ambiguïtés et en rendant les décisions plus uniformes.

5.1.4 ANALYSE DU COMPORTEMENT DU SYSTÈME

Lors du développement d'un système d'aide à la décision, différents besoins et contraintes sont pris en compte et intégrés dans le produit final. Toutefois, le volume important de connaissances (savoir et savoir-faire) implantées dans un tel système peut faire en sorte qu'il devient très difficile de s'assurer de la qualité des recommandations proposées. L'ajout d'explication sur le raisonnement du système peut avoir certains avantages au niveau de l'analyse de ses comportements et par le fait même de sa justesse dans tous les cas. Les éléments suivants présentent certains avantages identifiés dans la littérature ou issus des expériences de Novasys.

5.1.4.1 Justification du traitement du système

Les explications servent à donner les raisons du comportement du système d'aide à la décision [Herrmann et al., 98]. Elles présentent les différents éléments d'information disponibles, mettent en évidence les relations entre ces éléments et décrivent le raisonnement du système en fonction de ces éléments et relations. Cette justification amène un avantage important pour l'analyse de la qualité du comportement du système.

5.1.4.2 Meilleure détection d'erreurs dans les connaissances et les stratégies de raisonnement

Les explications sont très avantageuses pour la mise au point du système, pour la détection d'erreurs dans les connaissances et les stratégies de raisonnement [Haton et al., 91]. Elle donne une certaine forme de transparence au système [Herrmann et al., 98]. Les explications permettent aux utilisateurs de réaliser que les avis du système d'aide ne sont pas nécessairement corrects [Swartout et Moore, 93]. De plus, avec les explications, il devient plus facile de prédire et de tester les effets de changements sur les opérations du système [Waterman, 86].

5.1.4.3 Facilite l'entretien du système

Selon [Chevrier et Haton, 91] in [Haton et al., 91], les explications permettent de rendre compte du comportement du système et de l'enchaînement de ses raisonnements lors de la résolution d'un problème de façon à permettre une amélioration des bases de connaissances ou des mécanismes de raisonnement. Pour [Wick et Thompson, 92], les explications sont très appropriées pour l'entretien des systèmes. Grâce à elles, le système est plus facile à corriger et par le fait même, il peut devenir plus rapide à développer [Waterman, 86].

5.1.5 GESTION DES CONNAISSANCES

Les entreprises accumulent avec les années un bagage important de connaissances qui ont été plus ou moins formalisées selon leur nature, les besoins et les contraintes de l'entreprise. Une partie très importante de ces connaissances sont tacites, sous forme de savoir-faire spécifique et accessibles seulement par les employés qui réalisent les tâches exigeant ces connaissances. Dans ce sens, la gestion des connaissances devient une préoccupation importante pour les entreprises. Toutefois, leur acquisition se limite souvent aux règles de manipulation des données nécessaires à la prise de décision. L'analyse du processus de décision s'en trouve alors réduite et les connaissances relatives à l'explication du raisonnement menant à la décision du système sont absentes et de ce fait non conservées pour d'éventuelle réutilisation. Les éléments suivants présentent certains avantages relatifs à l'implantation des techniques d'explication pour la gestion des connaissances de l'entreprise.

5.1.5.1 Augmentation du niveau de connaissances du domaine d'expertise pour l'entreprise (décision + explication)

Le développement d'un système d'aide à la décision exige l'acquisition et la modélisation d'un ensemble important de connaissances sur un domaine d'expertise spécifique. Toutefois, avec l'ajout des explications des comportements du système, d'autres connaissances doivent être acquises et formalisées. Ce travail supplémentaire fait en sorte d'augmenter le niveau de connaissances connues sur un domaine d'expertise de l'entreprise. Non seulement la connaissance d'un expert est formalisée, mais la

connaissance que cet expert a de ses propres connaissances (ses méta-connaissances) [Richard et al., 92].

5.1.5.2 Réutilisation des connaissances

Le fait de formaliser et d'intégrer des connaissances explicatives dans un système d'aide à la décision permet de conserver cette expertise spécifique et de la réutiliser à grande échelle dans le futur.

5.1.5.3 Outil pédagogique

Les explications dans un système d'aide à la décision sont vues comme un outil pédagogique de diffusion et de transfert des connaissances permettant à une personne de mieux s'approprier les méthodes de raisonnement et les comportements d'un certain domaine d'activité [Haton et al., 91].

5.2 ANALYSE GÉNÉRALE DES COÛTS DES TECHNIQUES D'EXPLICATION

L'analyse des coûts inhérents à l'implantation de techniques d'explication dans un système d'aide à la décision a été regroupée en six thèmes spécifiques de manière à simplifier leur présentation. Ces thèmes sont les suivants :

- Développement du système;
- Tests du système;
- Traitement du système;
- Aspects psychologiques;
- Gestion des connaissances;
- Environnement technologique.

Cette section permet de décrire chacun des éléments de coûts ou d'efforts identifiés dans la littérature et lors des différents projets réalisés chez Novasys.

5.2.1 DÉVELOPPEMENT DU SYSTÈME

L'implantation de techniques d'explication dans un système d'aide à la décision exige d'investir des efforts supplémentaires. Les éléments suivants décrivent certains coûts liés à la phase de développement des techniques d'explication.

5.2.1.1 Acquisition des connaissances explicatives

Le développement de fonctions explicatives dans un système d'aide à la décision nécessite des connaissances spécifiques en plus de celles utilisées pour le raisonnement du système [Wick et Thompson, 92]. Cette activité d'acquisition peut être plus ou moins importante selon la complexité et la disponibilité de l'expertise, l'ampleur des connaissances en jeu, les stratégies d'explications nécessaires, le mode d'interaction explicative, les besoins et connaissances des utilisateurs, etc. C'est pourquoi, le volume de connaissances nécessaires pour l'explication peut être largement supérieur au volume des connaissances utilisées pour le raisonnement.

5.2.1.2 Modélisation des connaissances explicatives

La modélisation des connaissances explicatives peut être plus ou moins exigeante selon les besoins d'explication des utilisateurs, la complexité et l'étendue du domaine d'expertise. En effet, selon le type d'interaction explicative (explication sur demande, graduelle, globale, avec exemples, etc.), les connaissances à modéliser seront différentes. De plus, un domaine d'expertise très complexe pourra exiger une modélisation des explications qui devra tenir compte des caractéristiques de l'utilisateur, du contexte, des différentes stratégies explicatives [Haton et al., 91] et des conditions d'utilisation de ces stratégies. Enfin, selon [Haton et al., 91], ces différents éléments font en sorte que la modélisation des explications nécessite une abstraction des connaissances plus élevée que pour le traitement du raisonnement du système d'aide à la décision.

5.2.1.3 Modélisation de l'interaction explicative

Comme nous l'avons souligné auparavant, l'ajout d'explication est de plus en plus jumelé avec le besoin d'interaction entre le système et l'utilisateur du système. Ceci est dans le but de s'assurer que ces explications sont bien comprises par l'utilisateur et qu'elles sont efficaces. La modélisation de l'interaction explicative a donc pour objectif de déterminer les différents scénarios (contexte) de demande d'explications, les connaissances explicatives impliquées et les moyens (stratégies, formes) utilisés pour présenter les explications. L'ampleur de cette activité de modélisation peut être très importante si les interactions sont nombreuses, si les utilisateurs ont des connaissances très différentes ou si le domaine d'expertise est très complexe.

5.2.1.4 Modélisation de l'audience (interlocuteurs)

À partir du moment où un système d'aide à la décision a la possibilité d'expliquer son raisonnement, de justifier ou de mettre en contexte ses comportements, les connaissances sur le profil des utilisateurs deviennent un besoin très important. Ce besoin est encore plus fort si le système offre la possibilité d'interagir avec lui. En effet, des stratégies et des formes distinctes pourraient être nécessaires pour assurer la compréhension des explications.

Selon [Swartout et Moore, 93], cette activité de modélisation de l'audience est très complexe et peut exiger beaucoup d'efforts. Elle peut demander de modéliser différentes connaissances de l'utilisateur comme son expertise, ses préférences (avec exemples multimédias, le niveau de détail), ses préoccupations, ses besoins d'information en fonction de ses responsabilités, le vocabulaire qu'il utilise et le type d'interaction souhaité.

Cette modélisation de l'interlocuteur doit être cohérente et cohésive avec les possibilités d'interaction du système (la modélisation des interactions) et les connaissances disponibles (la modélisation des connaissances explicatives).

5.2.1.5 Temps de développement du système plus important

Avec les activités d'acquisition et de modélisation présentées précédemment, il devient évident que l'ajout d'explication dans un système d'aide à la décision augmente la complexité et les efforts de développement du système. Plus le système offrira des possibilités d'explication et sera sensible au profil des utilisateurs, plus les coûts seront importants.

5.2.2 TESTS DU SYSTÈME

L'implantation de techniques d'explication dans un système d'aide à la décision ne se fait pas sans ajouter des activités spécifiques de tests des explications. Les éléments suivants présentent différents coûts relatifs aux tests des explications fournies par le système.

5.2.2.1 Travail de validation et de calibrage des explications

Un système d'aide à la décision qui intègre des explications exige de mettre des efforts pour sa validation et son calibrage en fonction de la solution proposée [Tou Ng et Mooney, 89] et des différents besoins et contraintes des utilisateurs. D'après [Swartout et Moore, 93], les principales qualités recherchées pour une explications sont les suivantes :

- Explication fidèle :

  L'explication doit être une représentation exacte du raisonnement du système. Une explication inexacte ou erronée est pire qu'aucune explication.

- Explication compréhensible :

  Il est clair qu'une explication est utile si elle est comprise. En effet, la terminologie utilisée dans l'explication doit être familière pour l'utilisateur. Le système doit être capable de proposer une explication qui prend en considération les connaissances de l'utilisateur, ses buts, ses préférences et ses intérêts. Le système doit être capable de fournir des explications à différents niveaux d'abstraction, de détails, selon différentes perspectives, en

fonction du langage de l'utilisateur. De plus, le système doit pouvoir s'assurer que ses explications soient comprises.

- Explication suffisante :

  Le système doit savoir de quoi il parle et doit fournir suffisamment d'explication à l'utilisateur. Il doit expliquer et justifier ses comportements, ses choix de stratégie d'explication. Il doit être en mesure d'expliquer le domaine de connaissances et définir les termes du domaine si certaines questions spécifiques lui sont demandées.

5.2.2.2 Volume des tests plus important

La mise en place de techniques d'explication augmente l'envergure d'un système d'aide à la décision. Étant donné que le volume des connaissances gérées par le système est plus important, le volume des tests du système est aussi plus élevé. De plus, des tests de cohérence entre les recommandations et les explications ainsi qu'entre les demandes d'explication (précision, complément d'information, etc.) et les réponses du système s'ajoutent. Toutefois, le fait d'avoir des explications peut simplifier la validation du raisonnement du système même si le volume des tests est plus important.

5.2.3 TRAITEMENT DU SYSTÈME

L'ajout de techniques d'explication automatique dans un système d'aide à la décision fait en sorte d'augmenter le volume de traitement des connaissances. En effet, le système ne se limite plus à poser des questions, analyser les connaissances et à recommander une solution, mais il doit aussi analyser son raisonnement, répondre aux questions de l'utilisateur et justifier son comportement. Les éléments suivants présentent différents coûts relatifs aux traitements et à la génération des explications du système.

5.2.3.1 Effets sur la performance du système (temps réponse diminué)

L'augmentation du volume de connaissances à traiter pour expliquer le raisonnement ou répondre à une interrogation de l'utilisateur fait en sorte d'augmenter le volume de traitement. C'est du moins ce qu'a souligné [Wick et Thompson, 92] pour le cas des systèmes experts avec explications. Dans le même sens, [Goguen et al., 83] précisent que non seulement la recommandation est générée, mais il peut être nécessaire de contruire et de diffuser les explications avec des justifications et des exemples [Goguen et al., 83]. C'est pourquoi, le traitement des explications peut être exigent et doit donc être analysé et conçu de manière à ne pas avoir trop d'effets négatifs sur l'efficience du système [Swartout et Moore, 93].

5.2.3.2 Traitement additionnel pour gérer les stratégies d'explication

Les systèmes d'aide à la décision intégrant des explications du raisonnement peuvent gérer différentes stratégies pour présenter ses explications. Les stratégies disponibles doivent être traitées en fonction du profil des utilisateurs, du contexte d'aide ou de leur efficacité à faire comprendre le raisonnement aux utilisateurs ou pour répondre à leurs questions. Ces traitements additionnels pour gérer les stratégies d'explication sont des coûts à considérer.

De plus, certains systèmes peuvent se retrouver avec différentes alternatives d'explication. Des efforts supplémentaires sont parfois nécessaires pour traiter ces alternatives afin d'éliminer les explications erronées ou inefficaces selon l'utilisateur [Goguen et al., 83].

5.2.3.3 Prolongation du temps d'utilisation du système

La lecture et la compréhension des explications générées avec la recommandation où les interactions avec le système en mode dialogue peuvent faire en sorte d'augmenter le temps d'utilisation du système. Toutefois, la qualité des décisions compense généralement ces efforts supplémentaires.

5.2.4 ASPECTS PSYCHOLOGIQUES

Le processus de prise de décision est souvent un processus complexe et les impacts d'une mauvaise décision sont parfois très importants et psychologiquement exigeants pour le décideur. Dans un tel contexte, les informations fournies sont capitales pour la réalisation de ses fonctions. Les éléments examinés dans cette sous-section proposent différents éléments psychologiques ayant un effet possible sur le coût de l'ajout d'explication dans un système d'aide à la décision.

5.2.4.1 Possibilité d'un manque d'attention de l'utilisateur face aux erreurs

La génération d'explication au besoin ou avec les recommandations finales pourrait laisser croire que le système contrôle très bien toutes les situations de support à la décision. Toutefois, des erreurs de comportement ou certaines limites au niveau des connaissances d'un système peuvent l'amener à proposer des recommandations erronées ou non optimales. Dans ces cas, l'utilisateur ne doit pas diminuer son niveau d'attention durant le raisonnement du système et toujours s'assurer de la qualité des résultats. Un système d'aide à la décision sans explication exige ou amène généralement à plus de vérification des résultats car peu d'information permet de valider ou de critiquer le raisonnement. Même si cette situation est peu probable, les explications pourraient rendre l'utilisateur moins attentif aux erreurs. Il serait donc intéressant que le système puisse s'assurer que ces explications soient comprises et exactes.

5.2.4.2 Problème de confiance si les explications ne sont pas bien liées avec les recommandations

Selon [Wick et Thompson, 92], un problème de confiance est envisageable si les explications ne sont pas bien liées avec les recommandations surtout dans les cas de système où la construction des explications est indépendante de la construction des recommandations. Ce genre de problème est coûteux si le système est en production et qu'il amène des erreurs d'interprétation des décisions car les utilisateurs perdront leur confiance envers le système. Cette perte de confiance peut être très difficile et longue à rétablir. La survie du système pourrait être remise en cause.

5.2.5 GESTION DES CONNAISSANCES

La gestion et la réutilisation des connaissances sont des aspects importants pour un grand nombre d'entreprises. Toutefois, ces activités exigent certains investissements. Les éléments suivants présentent certains coûts à considérer pour la gestion des connaissances lors de l'implantation de techniques d'explication dans un système d'aide à la décision.

5.2.5.1 Gestion de l'évolution des connaissances explicatives

L'ajout de connaissances explicatives amène inéluctablement une augmentation des efforts nécessaires pour entretenir ses connaissances [Wick et Thompson, 92] ; [Swartout et Moore, 93]. De plus, ce type de connaissances peut être plus complexe et dans une proportion plus importante que les connaissances utilisées pour le raisonnement du système d'aide à la décision.

5.2.5.2 Gestion des relations entre les connaissances nécessaires au raisonnement et à son explication

Pour [Wick et Thompson, 92], l'ajout de connaissances explicatives aux connaissances nécessaires pour le raisonnement du système d'aide à la décision exige des efforts pour gérer les relations entre ces deux groupes de connaissances distincts.

> Par exemple, au lieu de gérer seulement un Savoir-faire « *X* », il est nécessaire de gérer les connaissances explicatives comme la Justification « *A* », les Profils utilisateurs « *B* » et « *C* », la Stratégie d'explication « *D* » qui permet de déterminer le Profil utilisateur « *B* » ou « *C* » et les Définitions « *E* » *et* « *F* ».

De plus, d'autres connaissances peuvent être nécessaires pour assurer la cohérence et la cohésion entre ces groupes de connaissances.

> Par exemple, si le Savoir–faire = « *X* » et le Profil utilisateur = « *B* » alors la Définition = « *E* ».

5.2.6 ENVIRONNEMENT TECHNOLOGIQUE

L'ajout de techniques d'explication dans un système d'aide à la décision peut amener le besoin de mettre en place certains éléments supplémentaires dans l'environnement technologique d'un système d'aide à la décision. Les points suivants présentent quelques éléments de coûts relatifs à l'intégration de techniques d'explication dans un système d'aide à la décision.

5.2.6.1 Technologies nécessaires

La conception d'un système d'aide à la décision avec des fonctions explicatives peut exiger une technologie différente (progiciel, système d'exploitation, etc.).

5.2.6.2 Architecture requise

L'architecture requise pour gérer les connaissances explicatives peut s'avérer différente de celle utilisée pour les connaissances pour le raisonnement. Par exemple, les systèmes experts de deuxième génération proposent des architectures qui séparent ces deux types de connaissances et les traitements qu'elles nécessitent.

5.2.6.3 Entretien et évolution de l'environnement du système

L'ajout des techniques d'explication fait en sorte de rendre le système d'aide à la décision plus important au niveau des traitements à effectuer, mais potentiellement aussi du côté de ses besoins architecturaux et technologiques. Dans ces cas, les coûts relatifs à l'entretien de ces environnements devront être considérés dans l'évaluation des efforts et des budgets.

5.3 TABLEAU DE SYNTHÈSE DE L'ANALYSE DES AVANTAGES ET DES COÛTS

AVANTAGES	COÛTS

Qualité des recommandations :
• Meilleure compréhension des problèmes complexes
• Augmentation du niveau de compréhension des recommandations données à l'utilisateur :
  • Raisons d'un comportement, contexte, relations entre les éléments, chronologie des comportements, solutions rejetées, hypothèses vérifiées
• Diminution des ambiguïtés
• Explication adaptable selon les besoins et contraintes des utilisateurs :
  • Niveau de dialogue, connaissances de l'utilisateur, préférences, prédisposition, explication sur demande (en mode dialogue), explication graduelle, explication avec exemples, plusieurs stratégies explicatives, ajout d'éléments explicatifs multimédias (image, son, animation, etc.)
 	
Aspects psychologiques :
• Augmentation du niveau de confiance
• Meilleure perception du comportement d'aide et de coopération fournies par le système
• Possibilité de diminution de la perception de contrôle par le système (le système donne plus de moyens pour le critiquer)
 	
Processus de prise de décision :
• Possibilité de rétroaction suite aux explications (pour demander plus d'explications, éclaircir certains éléments, ajouter des éléments d'information, etc.)
• Réduction du temps alloué à la prise de décisions
• Augmentation de la qualité des décisions
 	
Analyse du comportement du système :
• Justification du traitement du système
• Meilleure détection d'erreurs dans les connaissances et les stratégies de raisonnement
• Facilite l'entretien du système
 	
Gestion des connaissances :
• Augmentation du niveau de connaissances du domaine d'expertise pour l'entreprise (décision + explication)
• Réutilisation des connaissances
• Outil pédagogique
	Développement du système :
	• Acquisition des connaissances explicatives
	• Modélisation des connaissances explicatives
	• Modélisation de l'interaction explicative
	• Modélisation de l'audience (interlocuteurs)
	• Temps de développement du système plus important
	
	Tests du système :
	• Travail de validation et de calibrage des explications
	• Volume des tests des explications plus important
	
	Traitement du système :
	• Effets sur la performance du système temps réponse diminué)
	• Traitement additionnel pour gérer les stratégies d'explication
	• Prolongation du temps d'utilisation du système
	
	Aspects psychologiques :
	• Possibilité d'un manque d'attention de l'utilisateur face aux erreurs
	• Problème de confiance si les explications ne sont pas liées avec les recommandations
	
	Gestion des connaissances :
	• Gestion de l'évolution des connaissances explicatives
	• Gestion des relations entre les connaissances nécessaires au raisonnement et à son explication
	
	Environnement technologique :
	• Technologies nécessaires
	• Architecture requise
	• Entretien et évolution de l'environnement du système

6. ANALYSE DÉTAILLÉE D'UNE ARCHITECTURE DE SYSTÈMES EXPLICATIFS

Les premières définitions des systèmes d'aide à la décision (SAD) concernent des systèmes destinés à aider les personnes ayant à prendre des décisions stratégiques (décideurs) dans des situations de décision mal structurées. Ils sont conçus de manière à faciliter le travail des décideurs et non à remplacer leur jugement. Généralement ces décisions sont d'un type qui ne peut être complètement supporté par des algorithmes. Contrairement à des systèmes experts qui concernent surtout des diagnostics, des méthodes d'inférence avec usage d'heuristiques, des décisions complexes et non structurées, les SAD visent des planifications stratégiques, des problèmes intégrés complexes, semi-structurés, des modèles de bases de données avec accès interactif. Les explications dans un tel système gagneraient de fait à être scindées en parties spécifiques comportant non seulement des connaissances sur les problèmes, mais encore sur comment les expliquer et à qui les expliquer.

L'objectif de ce chapitre est de faire une analyse détaillée des différents problèmes d'aide à la décision et de leurs besoins d'explication respectifs. Il présente également une architecture générique d'explications en mettant l'emphase sur ses avantages et ses limites. Cette architecture est basée sur un ensemble de composants dont nous détaillons les caractéristiques. Le prototype d'un système d'explication présenté au chapitre suivant est en fait une instanciation de cette architecture.

Enfin, ce chapitre indique comment le recours à une architecture basée sur des agents intelligents permet d'intégrer les techniques d'explications prometteuses.

6.1 DIFFÉRENTS PROBLÈMES D'AIDE À LA DÉCISION

Le processus d'explication d'un problème de décision nécessite une analyse préalable de la nature du problème. De là, il s'agit de déterminer quel type d'explication pourrait convenir à ce type de problèmes. Cette section aborde différents types de problèmes d'aide à la décision et associe des types d'explications appropriés à l'aide d'un tableau synthétique.

Type de problème	Description du problème	Type d'explication possible
Aide au diagnostic	Identifier des problèmes de fonctionnement à partir d'observations.	Les explications sont généralement des déductions faites à partir des symptômes vers les conclusions. Il peut s'agir de diagnostic médical, mécanique, électronique, etc.
Aide à la prédiction	Prédire des conséquences de situations données.	Les explications peuvent être à la fois des déductions de conséquences probables à partir de situations données ou des inductions visant à généraliser un phénomène. Par exemple, des prévisions militaires, météorologiques, démographiques, etc.
Aide à l'interprétation	Décrire des situations à partir de données sensorielles.	Il s'agit de donner les raisons de certaines configurations ou situations provenant de systèmes complexes comme la reconnaissance, la parole, l'analyse d'images ou de textes ou encore l'interprétation de signaux.
Aide à la planification	Déterminer des plans d'action.	Il s'agit d'expliquer l'ordre de la séquence d'actions dans différents domaines tels que la planification de projets, les étapes d'un robot ou l'acheminement de marchandises, etc.
Aide à la surveillance et au contrôle	Comparer des observations à des défaillances possibles et remédier à ces défaillances.	Il s'agit de veiller à expliquer les changements du comportement d'un système, comme le contrôle aérien, celui des missions, etc.
Aide à la réparation	Prescrire des remèdes à des mauvais fonctionnements.	Il s'agit ici de pouvoir expliquer le choix d'un remède plutôt qu'un autre afin de corriger le mauvais fonctionnement. Les différents critères peuvent être respectivement considérés pour justifier ce choix. Cela peut être le coût de la réparation, la vitesse, l'opportunité, les ressources impliquées, etc.
Aide à la formation	Identifier des failles dans les connaissances de l'usager et proposer des solutions pour y remédier.	Il s'agit de considérer les stratégies qui conviennent à l'apprenant en considérant non seulement son niveau de connaissances sur le sujet d'apprentissage, mais encore son style et ses critères affectifs. Le rythme des explications et leur granularité sont fonction de ces critères, l'important étant de présenter la bonne explication au bon moment, de façon graduelle ou globale.
Aide à l'évaluation et à la sélection d'options	Identifier la ou les meilleures solutions possibles en fonction des données du problème (faits, observations, etc.) et en fonction des contraintes et objectifs ajoutés pour orienter la sélection de la solution	Il s'aggit d'expliquer comment les contraintes et les objectifs définies pour la recherche de la meilleure solution ont influencés le choix de la solution et le rejet des autres. Le système peut aussi expliquer les effets de différents changements de contraintes et d'objectifs sur la solution retenue. Dans ce sens, il peut servir pour simuler différents scénarios.

6.2 REPRÉSENTATION DES CONNAISSANCES

6.2.1 DIFFÉRENTS NIVEAUX DE CONNAISSANCE

Parmi des techniques relativement récentes, l'article de Jurgen Herrmann et al [J. Herrmann, M. Kloth, Frank Feldkamp, 1998] présente une approche d'explications basée sur quatre couches qui assurent une variété d'explications pour des systèmes d'aide à la décision. Elle offre notamment une distinction entre les parties statiques et dynamiques.

Les parties de systèmes dynamiques concernent des systèmes dont les solutions dépendent du temps (résolution de problèmes) alors que les parties statiques concernent les caractéristiques générales du système. Les quatre niveaux concernent le *domaine*, le *système*, le *processus* et le niveau *fonction*.

- Le niveau « domaine » comporte toutes les connaissances utiles à l'explication et relatives au domaine (objets, relations, actions typiques).
- Le niveau « système » permet à l'usager de comprendre comment le système est intégré dans le domaine et quel sont ses capacités et ses limitations au niveau des type de problème, des actions et des stratégies qu'il peut traiter dans le domaine.
- Le niveau « processus » contient la connaissance relative aux explications qui est générée et manipulée pendant l'interaction avec le système.
- Le niveau « fonctionnel » contient les différentes fonctions fournies pour le système de résolution de problèmes.

Si cette approche a l'avantage de séparer les connaissances et de fournir plus de flexibilité que les autres systèmes d'explication, elle ne considère pas comme fondamental le niveau de perception de l'usager. Pour une architecture générale nous proposons trois groupes de connaissances:

Groupe 1 : *Connaissances* permettant de manipuler des données, comme celles relatives à la résolution de problèmes et à l'explication.

Groupe 2 : *Faits ou données*, comme la base des problèmes, la base de l'historique des solutions, la base de l'historique des explications, la base de profils utilisateurs et la base terminologique.

groupe 3 : *Métaconnaissances* comme celles relatives à la justification des connaissances sur les méthodes explicatives ou de résolution de problèmes.

Ces groupes sont détaillés à la section 6.2.3.

6.2.2 ARCHITECTURE PROPOSÉE

6.2.2.1 L'architecture du système d'explication

L'architecture proposée dans cette étude et présentée à la figure 8 contient deux systèmes : le *générateur de solution* et le *générateur d'explications*.

Le générateur de solutions a pour rôle de produire une solution à partir des données du problème (1). Ces données sont extraites de différentes sources (0). (2) est l'élément central de ce traitement car elle permet de s'assurer que les connaissances sont suffisante pour prendre une décision et, dans le cas contraire, diriger dans les demandes de questions supplémentaires (4). De plus, la base de connaissances sur les méthodes de résolution de problèmes permet de tirer partie de la réutilisation potentielle des solutions antérieures contenues dans la base historique des solutions (3). Les réponses aux questions sont conservées dans le fichier des données du problème (5). La solution (6) est transmise à l'usager et stockée dans une base de solutions en vue de réutilisations futures et d'utilisation par le système d'explication.

À la réception de la solution du problème courant (9), le générateur d'explications utilise la base de connaissances sur les méthodes explicatives (7) et la base historique des explications (13) pour diriger son traitement. Il utilise les données du problème incluant les réponses aux questions supplémentaires (8), les connaissances sur les méthodes de résolution de problèmes (10), les connaissances terminologiques (12) et la base de profils utilisateurs (11) pour adapter un premier niveau d'explication. Dans le cas d'explications plus détaillées (par exemple, pour la validation du bon fonctionnement du système), le générateur utilise les connaissances sur les justifications des méthodes de résolution des problèmes (16) et d'explications des solutions (17) pour produire ces explications.

Enfin, durant le traitement de résolution d'un problème, il est possible que l'usager ait besoin d'explications (15). Dans ce cas de dialogue, le système d'explication se chargera de répondre à sa demande (14) en utilisant les connaissances disponibles.

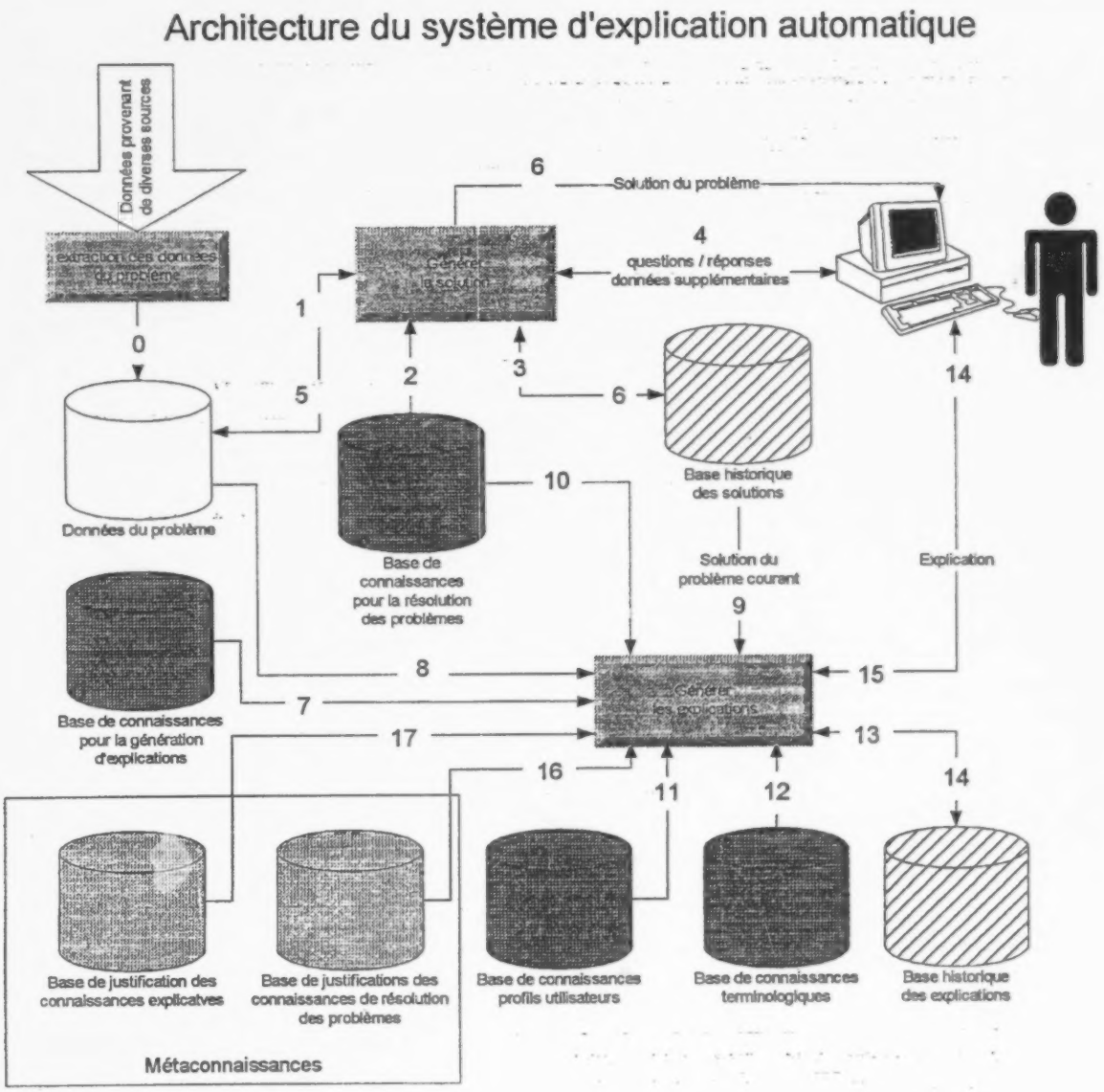

Figure 8. Architecture du système d'explication automatique

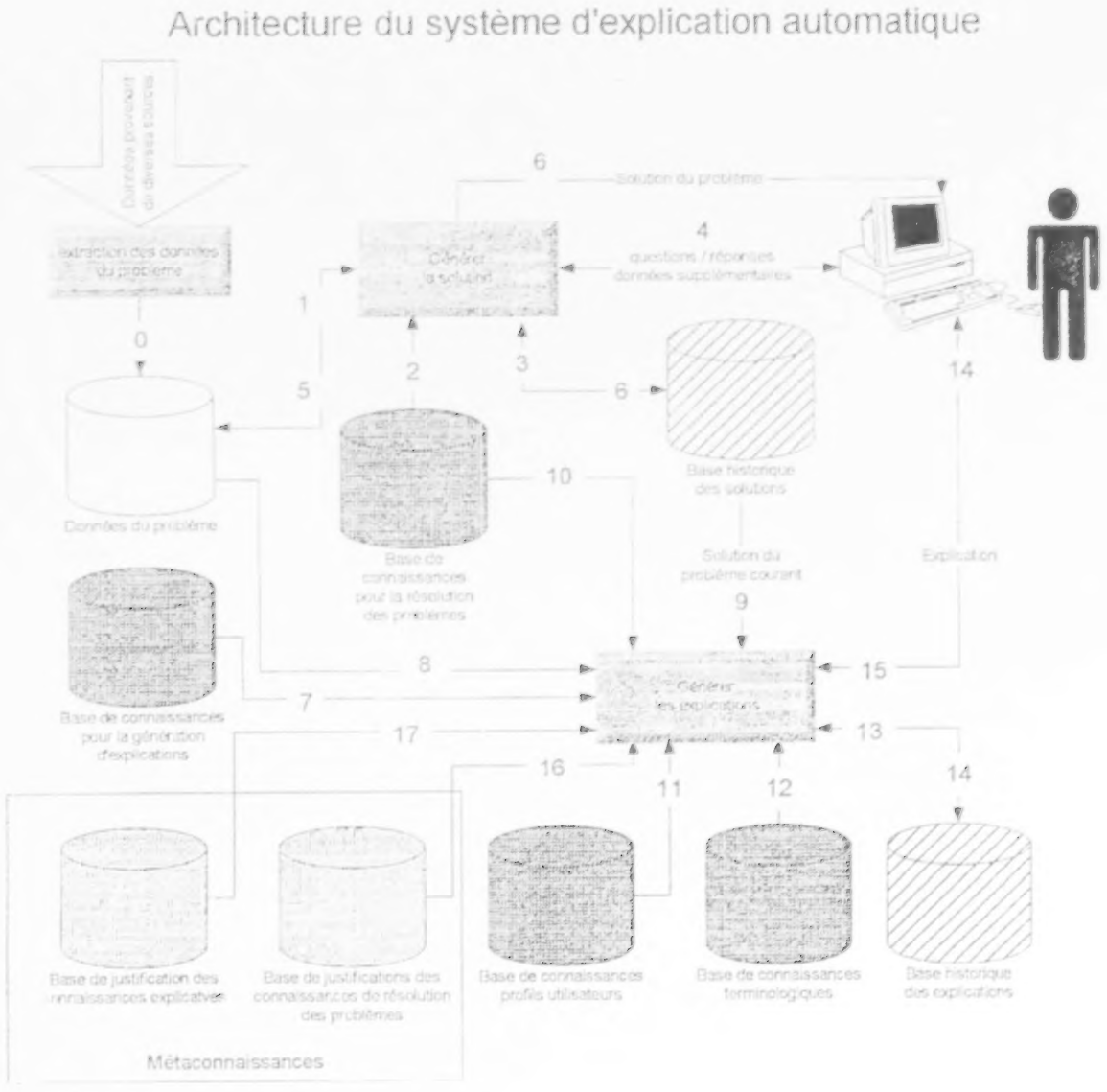

Figure 8. Architecture du système d'explication automatique

6.2.2.2 Avantages de l'architecture

Modularité : cette architecture décompose les connaissances de façon modulaire, ce qui assure une distribution spécifique des connaissances et permet une mise à jour plus facile de chaque module.

Niveau de métaconnaissance : ce niveau permet de déployer effectivement les explications selon les besoins des usagers. En fonction du profil de l'usager, le système sera à même d'adapter les explications à la fois selon le sujet et le niveau de l'usager.

Architecture ouverte : il est possible de rajouter d'autres modules selon le besoin de raffinement nécessaire. Ainsi, il serait possible de rajouter des explications spécifiques à un domaine ou des contextes particuliers.

6.2.2.3 Limites de l'architecture

Augmentation des échanges : la multiplicité des bases implique un plus grand nombre de communications avec le système de génération d'explications. Ceci entraîne des protocoles de communication plus complexes.

Difficulté de constitution : étant donné que chaque base comprend des expertises spécifiques, il peut être plus long de les acquérir et de les modéliser. De plus, il faut veiller à bien les indexer afin de retrouver le plus rapidement possible les informations souhaitées. Par exemple, on pourra indexer selon le type de problèmes (diagnostic de violation de l'espace aérien), sa difficulté (avion armé ou non), sa complexité (causes de violation).

Validation des connaissances : il s'agit de savoir si les bases de connaissances de départ sont fiables et ensuite de vérifier que les résultats (explications données) ne contredisent pas les données de départ et sont en accord avec des spécialistes du domaine.

6.2.3 DÉTAILS DE L'ARCHITECTURE DES BASES DE CONNAISSANCES

L'architecture des connaissances est articulée autour de trois groupes de connaissances.

6.2.3.1 Les connaissances sur les méthodes

Ce groupe de connaissances sert à diriger les traitements de résolution de problème et d'explication des solutions et dicte comment utiliser les autres connaissances disponibles.

Base des connaissances de résolutions de problèmes

Elle comporte les formalismes de résolution de problèmes comme des règles, des objets, des réseaux sémantiques, des grilles d'analyse, etc. Ce sont des connaissances qui permettent au système de savoir comment résoudre les problèmes.

Elle comporte aussi les stratégies de résolution selon une classe de problème On ne procédera pas de la même manière pour résoudre un problème de diagnostic ou de planification.

Base des connaissances explicatives

Elle comporte des connaissances sur comment expliquer la résolution d'un problème. Ces connaissances sont représentées sous forme de règles, d'objets, de réseaux sémantiques, de grilles d'analyse, des gabarit (des « frames » ou « templates » selon [Herrmann et al. 98]). On a ici des stratégies d'explications selon le profil : explications ponctuelles (données de temps en temps), graduelles (données étape par étape), globales (données à la fin), référencées (faisant référence à d'anciennes explications, avec exemples, etc.) ainsi que sa connaissance plus ou moins approfondie du domaine d'expertise du système et des mécanismes automatiques de résolution de problèmes.

6.2.3.2 Les connaissances complémentaires aux méthodes

Ce groupe de connaissances est utilisé durant le traitement des méthodes de résolution et d'explication. Ces connaissances fournissent le contenu nécessaire à la résolution et à l'explication d'un problème.

Données du problème

Ce fichier contient des énoncés de problèmes sur lesquels des solutions n'ont pas été apportées et que l'on souhaite soumettre à l'utilisateur. Il contient aussi les réponses aux questions supplémentaires demandées en cours de résolution.

Base de l'historique des solutions

Elle contient un ensemble de solutions relatives aux problèmes résolus. L'indexation des problèmes permet de retrouver les solutions à partir des énoncés de la base de problèmes

résolus. Les résultats de la génération de solutions sont placés au fur et à mesure dans cette base.

Base de l'historique des explications

Elle contient les explications relatives aux problèmes résolus. Il s'agit d'explications qui ont été générées pour clarifier les solutions des problèmes résolus et qui sont placées successivement dans cette base. Il est possible d'indexer une explication à partir du problème ou de sa solution.

Base des profils utilisateurs

Un utilisateur peut être décrit selon trois approches : cognitive, affective ou inférentielle. La partie cognitive traduit l'état des connaissances de l'utilisateur par rapport au problème posé (nul, moyen ou fort). La partie affective concerne les préférences de l'utilisateur quant à une stratégie d'explication donnée ou un mode d'apprentissage habituellement utilisé. La partie inférentielle concerne son mode de raisonnement (déduction, induction, analogie). La base contient également l'historique des problèmes résolus pour chaque utilisateur.

Base de connaissances terminologiques

Elle comporte la description des termes utilisés dans les questions et des précisions pour des explications. De plus elle peut contenir des contraintes sur certaines connaissances. Exemple : quand les utiliser, avec quelles connaissances il ne faut pas les utiliser etc.

Elle comporte enfin tous les termes du dictionnaire qui sont utilisés dans les énoncés de problèmes, les solutions ou les explications y compris les synonymes.

6.2.3.3 Les métaconnaissances

Ce type de base de connaissances sert à conserver des connaissances sur les raisons et la nature des méthodes de résolution et d'explication de la solution. Ces connaissances sur les connaissances relatives aux méthodes sont considérées comme des « métaconnaissances ».

Base de justifications des connaissances de résolutions de problèmes

Elle comporte des métaconnaissances sur les connaissances de résolution de problèmes. Différents facteurs peuvent être pris en compte comme la date, la source (un ou plusieurs experts), le contexte d'utilisation (culturel, sociologique, etc.), des connaissances de bon sens. Elle peut contenir aussi des heuristiques que l'expert utilise pour pouvoir arriver le plus vite possible à la solution du problème.

Base de justifications des connaissances explicatives

Il s'agit de métaconnaissances sur les connaissances explicatives. Différents facteurs peuvent aussi être pris en compte comme la date, la source (un ou plusieurs experts), le contexte d'utilisation (culturel, sociologique, etc.). Elles traduisent des points de vue de différents experts et donc des heuristiques différentes aussi.

6.2.4 FORMALISME ET REPRÉSENTATION DES CONNAISSANCES

Pour résoudre ces problèmes d'aide, les connaissances doivent être représentées selon un formalisme donné. Cela peut être: des règles, des objets, des « frames », des réseaux sémantiques, des grilles d'analyse, etc. Pour représenter les connaissances (de résolution ou d'explication), on peut choisir le même formalisme ou des formalismes différents. On peut même combiner différents types de formalismes.

6.3 TRAITEMENTS SUR LES CONNAISSANCES

6.3.1 INTRODUCTION

On peut distinguer deux types de traitements supportés par l'architecture générale proposée : la génération de solution et la génération d'explication. Nous ne nous intéresserons pas à la première puisque cette solution est générée par un système existant actuellement. La génération d'explications est un processus plus complexe qui, comme l'indique la figure 8, utilise de nombreuses sources de connaissances. La multiplicité de ces sources lui permet d'être flexible et surtout de fournir une variété d'explications selon les usagers.

Avant de générer les explications, le système est à même de répondre à une question de l'utilisateur pour préciser un terme ou encore les données du problème. La génération des explications est réalisée en fonction des connaissances sur :

- les méthodes de résolution du problème;
- les justifications des méthodes de résolution des problèmes;
- les méthodes d'explication des solutions;
- les justifications des méthodes d'explication;
- les données du problème;
- le profil d'utilisateur;
- les connaissances terminologiques.

Enfin, selon [Herrmann et al., 98], les principaux traitements relatifs à l'explication d'une solution sont :

- la collecte des informations nécessaires à l'explication;
- la structuration de l'explication;
- l'adaptation de l'explication en fonction de l'auditoire et;
- la présentation de l'explication à l'auditoire.

6.3.2 UNE ARCHITECTURE DE TRAITEMENT PROMETTEUSE BASÉE SUR UNE APPROCHE PAR AGENTS INTELLIGENTS

6.3.2.1 Définitions

Les agents intelligents constituent à notre avis une approche prometteuse pour le type de problème qui est à traiter. Avant de le montrer, nous allons préciser ce que nous entendons par agents.

La définition des agents intelligents varie selon les auteurs, mais nous ne retiendrons que quelques-unes d'entre elles. Selon Pattie Maes [Maes, 1995], un agent autonome est un système électronique qui habite un environnement complexe, qui perçoit son environnement et agit à l'intérieur de celui-ci de façon autonome pour réaliser un ensemble de buts ou de tâches pour lesquels il est conçu.

Selon Tecuci [Tecuci, 1998], un agent autonome est un système à base de connaissances qui peut interagir avec un usager (ou d'autres agents). Il est nécessaire, toutefois, d'appliquer des formes d'apprentissage faisant appel à plusieurs stratégies pour que les explications soient efficaces.

Selon Hayes-Roth [Hayes-Roth, 1995] un agent intelligent effectue continuellement trois fonctions : percevoir les états dynamiques de leur environnement, agir sur ces états dynamiques et raisonner afin d'interpréter les perceptions, d'effectuer des inférences, de résoudre des problèmes et de déterminer les actions à entreprendre.

Selon Frasson [Frasson, 1998], les agents sont réactifs, planificateurs et cognitifs, c'est-à-dire qu'ils ont des capacités d'apprentissage sur des nouvelles situations et décisions à prendre dans ces cas.

Les principaux attributs d'un agent sont :

- L'autonomie (opération sans intervention directe des humains);
- La réactivité : les agents perçoivent leur environnement et agissent en temps opportun aux changements qui se produisent;
- L'initiative : les agents sont capables d'exhiber des comportements commandés par des buts en prenant l'initiative.

Les architectures des agents sont souvent constituées de couches qui isolent les connaissances à mettre en œuvre. Elles permettent aussi de constituer des réseaux d'hypothèses sur des conduites à tenir dans des situations diverses. Ainsi, l'architecture modulaire que nous proposons peut constituer une base de réalisation d'agents dans la mesure où les connaissances sont spécialisées et réparties. La communication entre agents peut se faire au moment de la présentation du cas à expliquer. La spécialisation des explications est surtout basée sur le profil des usagers qui peut être pris en charge par

un agent spécialisé. Cette approche va plus loin que l'architecture en couche proposée par [Herrmann et al., 98] dans la mesure où chaque couche constitue un agent autonome.

Une approche des agents souvent reconnue consiste à les décrire selon trois modèles. Un modèle de *croyances* qui décrit les informations sur l'environnement et les états internes qu'un agent peut gérer et les actions qu'il peut réaliser. Ces croyances sont décrites selon des ensembles d'états. Un modèle de *buts* qui décrit les buts qu'un agent peut vraisemblablement adopter et les événements qui y correspondent. Un modèle de planification qui décrit les plans qu'un agent peut éventuellement utiliser pour réaliser ses objectifs et répondre aux évènements qu'il perçoit. Dans le contexte actuel des explications, les caractéristiques des agents peuvent être décrites selon des situations intervenant dans l'espace aérien pour lesquels un certain nombre de croyances sont établies à la fois concernant le problème visé et les usagers considérés. Selon ces deux derniers paramètres, l'explication peut être adaptée et détaillée davantage. Dans cette approche il est possible de mieux considérer les aspects relevant des objectifs de chaque agent, par exemple les objectifs des explications, celles concernant les justifications (connaissances explicatives et connaissances de résolution de problèmes). L'adaptation des explications peut tenir compte des croyances quant au niveau de connaissance de l'usager et aux situations d'incompréhension possibles.

6.3.2.2 Avantages de l'approche

Les agents commencent à être utilisés dans des domaines complexes comme le contrôle de trafic aérien, la gestion des ressources du personnel aérien, la gestion de l'espace aérien, les systèmes de simulation, les processus d'affaires et la gestion des échanges financiers, la gestion des réseaux de communication.

Les avantages des agents sont de pouvoir constituer des systèmes qui s'améliorent avec le temps et qui sont capables de considérer, de façon autonome, des variétés de situations selon des buts déterminés. Ils sont d'une aide prometteuse dans les SAD car ils peuvent se spécialiser selon les types de connaissances et donc affiner leurs recommandations. Dans le cadre de systèmes de formation à la décision, la séparation entre agents spécialisés permet d'améliorer au fur et à mesure les connaissances de chacune des bases sur lesquelles les agents interviennent. Contrairement aux simples systèmes de bases de connaissances et de tableaux noirs, l'approche avec agent permet d'avoir des entités autonomes spécialisées dans chacune des connaissances mises en œuvre lors des explications et d'agir selon des hypothèses et des buts dans des situations spécifiques.

Il serait possible d'aborder la définition et composition d'une architecture basée sur les agents ainsi qu'une implantation, dans une étude plus large.

6.3.2.3 Quelques exemples d'agents pouvant être intégrés à l'architecture générale d'un système d'aide à l'explication des décisions

Plusieurs agents spécialisés pourraient être développés pour répondre à des besoins spécifiques relatifs aux explications données par un système d'aide à la prise de décisions. Voici quelques exemples d'agents applicables à l'architecture générale proposée et les bases de connaissances utilisées pour leurs traitements :

	Agents proposés	Bases de connaissances associées								
		Base des connaissances de résolutions de problèmes	Base des connaissances explicatives	Données du problème	Base de l'historique des solutions	Base de l'historique des explications	Base des profils utilisateurs	Base de connaissances terminologiques	Base de justifications des connaissances de résolutions de problèmes	Base de justifications des connaissances explicatives
1	Collecter les connaissances pour les besoins d'explication	X	X	X	X	X	X	X	X	X
2	Structurer les connaissances pour les besoins d'explication	X	X	X	X	X	X	X	X	X
3	Adapter les connaissances à l'auditoire		X	X		X	X	X	X	X
4	Présenter l'explication en fonction des médias utilisés ou disponibles		X			X	X			
5	Répondre à une demande d'explication	X	X	X	X	X	X	X	X	X
6	Répondre à une demande de précision d'une explication ou d'un terme spécifique		X	X		X	X	X	X	X
7	Gérer le niveau de détail des explications		X				X			
8	Répondre à une demande de justification d'une explication		X			X	X		X	X
9	S'assurer que l'utilisateur a bien compris une explication		X	X		X	X	X		
10	Analyser les demandes d'explications et proposer des adaptations aux explications		X	X	X	X	X	X	X	X
11	Proposer des explications alternatives basées sur des stratégies différentes		X			X	X			
12	Analyser les situations d'explication antérieures et proposer des explications basées sur ces connaissances		X			X				
13	Coordonner et contrôler le travail des autres agents (la société d'agent)									
14	Analyser les actions des utilisateurs à la suite des recommandations du système									
15	Echanger des connaissances avec d'autres systèmes									

6.4 INTERFACES

Les caractéristiques les plus importantes des interfaces dans un SAD sont les suivantes.

Il existe deux types d'interface, chacune permettant une interaction avec un type d'opérateur. Ainsi, nous considérons l'interface avec les *experts* et les cogniticiens pour construire la base de connaissances, et l'interface avec les *usagers* qui ont recours au SAD.

L'interface avec les experts nécessite des outils de communication permettant des dialogues, des interviews, des analyses de discussions. Cette interface ne constitue pas l'objet de la présente étude.

L'interface avec les usagers peut prendre plusieurs formes :

- Par interaction avec des menus dans lesquels l'usager choisit un item parmi une liste de choix possibles selon la fonction à accomplir. Les menus apparaissent selon un ordre logique, pouvant aller de menus en sous-menus. Cette approche peut toutefois devenir ennuyeuse dans le cas de situations complexes à analyser.

- Les langages de commandes utilisent des mot-clés qui déclenchent des macros ou procédures spécifiques.

- Le remplissage de formulaires est un moyen de rentrer des données dans des champs spécifiques, proposés par le système au fur et à mesure de l'interaction.

- Le langage naturel reste un moyen complexe car les moyens d'interprétation sont encore complexes à mettre en œuvre et les résultats sont limités. Il est difficile de considérer de tels systèmes dans le cas de commandes et réponses rapides.

- Les interfaces graphiques sont souvent représentées par des icônes qui sont directement manipulées par l'usager. Elles peuvent s'avérer longues à utiliser.

Il est à signaler que les commandes vocales constituent actuellement des moyens d'interaction rapides particulièrement intéressants pour des SAD visant des décisions militaires. Des commandes vocales telles que pourquoi ou comment peuvent alors déclencher les explications selon le contexte de l'interaction.

6.5 CONCLUSION

Ce chapitre a permis de présenter un certain nombre de détails concernant les explications d'un système d'aide à la décision. Il a porté davantage attention par l'étude préalable concernant l'ensemble des techniques relatives au problème considéré. L'architecture que nous proposons résulte à la fois des tendances récentes dans les solutions relatives aux systèmes d'explications et aux moyens susceptibles de les mettre en œuvre. Ainsi, la séparation en plusieurs niveaux de connaissances spécifiques permet une plus grande flexibilité d'adaptations des explications aux situations potentielles.

La technologie des agents est une voie prometteuse qui permettrait d'implanter l'approche retenue. Elle procure des moyens beaucoup plus performants que les approches conventionnelles et assure la mise en œuvre d'entités autonomes permettant de raisonner selon des situations et connaissances spécifiques. Cette tendance se confirme d'ailleurs par le recours récent à ces techniques pour des problèmes similaires.

Nous recommandons donc cette approche qui pourrait faire l'objet d'une étude spécifique avec implantation. Dans le but de fournir une idée des fonctionnalités permises par l'architecture, nous allons présenter un prototype appliqué à un exemple concret de la problématique des explications dans le domaine de la gestion de l'espace aérien.

7. PROTOTYPE D'UN SYSTÈME EXPLICATIF

Ce chapitre présente le prototype (maquette) d'un système d'explications automatique. Il est inspiré de l'architecture générale détaillée au chapitre précédent. Toutefois, ce prototype n'intègre pas l'ensemble des concepts et fonctions proposés dans l'architecture générale, ceci en raison des délais et du temps accordés au développement de ce dernier. Néanmoins, il exploite les principales idées apportées par l'analyse et sert de tremplin pour un projet d'une envergure beaucoup plus grande.

Trois sections distinctes divisent ce document. Les principales caractéristiques techniques rattachées au prototype sont présentées en premier lieu. Ensuite, le mode d'installation propre au bon fonctionnement du prototype et les principales options offertes par ce dernier sont abordés. Enfin, les connaissances relatives aux différents paramètres du prototype sont exposées.

7.1 POINTS TECHNIQUES

Le langage de programmation utilisé au développement du prototype est JAVA ; plus précisément, nous avons utilisé l'ensemble JDK1.1.7A et la librairie graphique Swing 1.0.3. Le tout a été développé et testé sous la plate-forme WindowsNT 4.0 et Windows 98, mais devrait être opérationnel sur les stations SUN ou Macintosh étant donné la portabilité de JAVA.

7.2 MODE D'INSTALLATION

Le prototype est composé de trois principaux fichiers, soit :

- JDK1.1.7A.exe
- Swing-1.0.3.zip
- Sources.zip

Tout d'abord, ce chapitre est initié par une présentation des principales étapes d'installation de l'ensemble de développement JDK1.1.7A et de la librairie graphique Swing1.0.3. Cette installation est fait sur l'ordinateur où résidera le prototype. Ensuite, il est question de l'installation du prototype lui-même.

7.2.1 INSTALLATION DE JDK1.1.7A

Pour lancer l'installation de JDK1.1.7A, il suffit d'appuyer sur l'icône de ce dernier avec la souris. Une suite d'instructions guide le suite de l'installation. De là, modifiez soit dans le fichier autoexec.bat pour Windows 98 ou soit les variables d'environnements pour Windows NT, les valeurs associées au *PATH* et au *CLASSPATH*.

- Ajoutez au *PATH* c:\jdk1.1.7A\bin\;
- Ajoutez au *CLASSPATH* c:\jdk1.1.7A\lib\classes.zip.

http://www.sun.com/solaris/jdk/download.1.1.7/

7.2.2 INSTALLATION DE SWING 1.0.3

Pour ce qui est de la librairie graphique Swing1.0.3, décompressez le fichier à la source de votre disque dur (ex : c:\). Par la suite, ajoutez à la variable *CLASSPATH* la valeur indiquée ci-dessous :

- Ajoutez au CLASSPATH c:\swing-1.0.3\swingall.jar;

Il est important de noter que les répertoires c:\jdk1.1.7A et c:\swing-1.0.3 peuvent être différents selon votre installation.

http://java.sun.com/products/jfc/index.html#download-1.0.3

7.2.3 INSTALLATION DU PROTOTYPE

L'installation du prototype consiste en une seule étape, celle de décompresser le fichier sources.zip dans un répertoire de votre choix. Afin de démarrer l'application, allez dans le répertoire choisi et appuyez deux fois à l'aide de votre souris sur le fichier dss.bat ; l'application devrait ainsi démarrer.

7.3 PROCESSUS GLOBAL DU PROTOTYPE

Le processus d'exécution général du prototype est illustré à la figure 9. Le point important du schéma est relié au fait qu'il est possible de mettre à jour les évènements (paramètre de la mission) après avoir consulté une première fois la décision suivie des explications. Il en est de même avec les objectifs de la mission, c'est-à-dire qu'il est possible de modifier les poids d'une mission et ensuite de visualiser la nouvelle décision suivie des nouvelles explications.

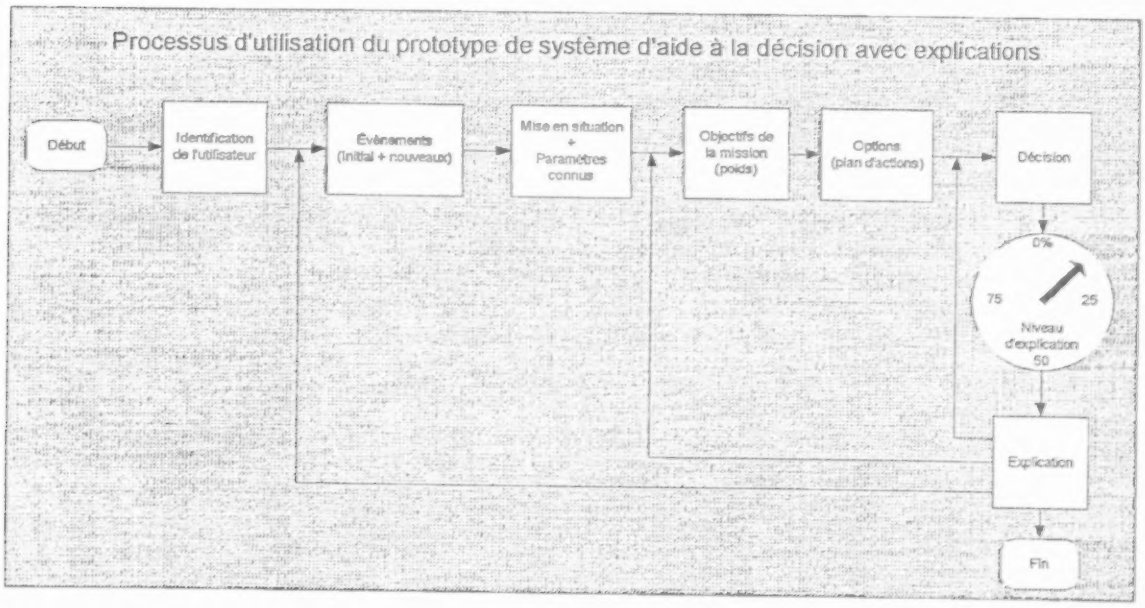

Figure 9 – Processus d'utilisation du prototype

7.3 PROCESSUS GLOBAL DU PROTOTYPE

Le processus d'exécution général du prototype est illustré à la figure 9. Le point important du schéma est relié au fait qu'il est possible de mettre à jour les évènements (paramètre de la mission) après avoir consulté une première fois la décision suivie des explications. Il en est de même avec les objectifs de la mission, c'est-à-dire qu'il est possible de modifier les poids d'une mission et ensuite de visualiser la nouvelle décision suivie des nouvelles explications.

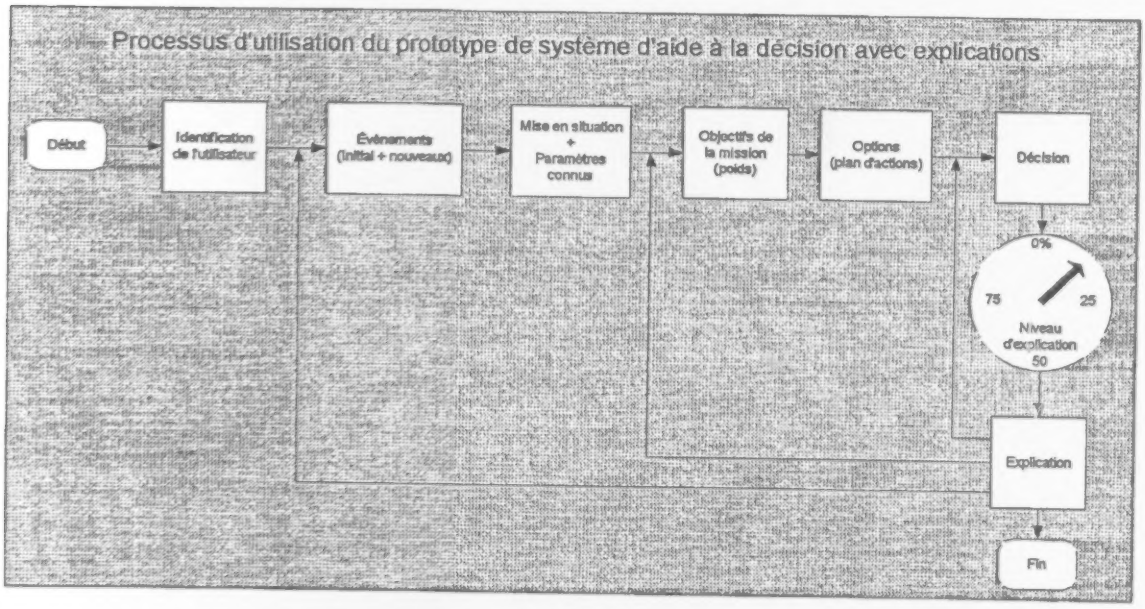

Figure 9 – Processus d'utilisation du prototype

7.4 PRINCIPALES OPTIONS

Cette section débute par un bref aperçu de la fonctionnalité première du prototype. Ensuite, elle présente chaque étape importante d'une séance d'utilisation du prototype. Il est important de mantionner qu'aucun agent intelligent n'a été développé dans ce prototype. Il ne permet que de faire une simulation d'un outil intelligent et présente une façon de présenter les interactions avec l'utilisateur.

7.4.1 VUE GLOBALE DU SYSTÈME

Ce prototype est un système d'aide à la décision dans lequel des explications sont générées selon les décisions. Ce système a pour but d'apporter une aide plus particulière à la planification, à la surveillance et au contrôle. Le domaine choisi porte sur la violation d'un espace aérien.

Ainsi, en fonction d'une situation initiale captée par divers types de senseurs, le système a pour rôle de proposer une solution à partir des données du problème et des différentes connaissances relatives à la résolution des problèmes. La solution est transmise à l'usager comme plusieurs options possibles. Après avoir pris en considération tous les paramètres de la situation, l'usager oriente le choix d'une option particulière en précisant le niveau de complexité et de risque envisagé pour la mission. La solution « idéale » est alors transmise à l'usager et stockée dans une base de solutions en vue de réutilisations futures.

On retrouve une ou plusieurs explications justifiant le choix de la solution. Le générateur d'explications utilise la base de connaissances explicatives (pour le traitement des explications selon les méthodes et connaissances du problème), les connaissances de résolution de problèmes (pour résoudre le problème), les connaissances terminologiques (pour préciser certains termes techniques) et la base de profils utilisateurs (pour adapter un premier niveau d'explication en fonction des préférences des utilisateurs). Il est important de noter que dans cette maquette, toutes les bases de connaissances sont statiques et codifiées directement. Donc, aucune intelligence n'est intégrée.

Dans une version réelle future, il serait important de permettre à l'utilisateur de demander des explications supplémentaires. Dans ce cas, le générateur d'explications devrait utiliser les justifications de connaissances explicatives et de résolution de problèmes pour produire ces explications.

Toutes les bases de données ou de connaissances citées ci-dessus sont décrites plus en détail à la section 6.2.2 (figure 8) de l'analyse détaillée.

7.4.2 IDENTIFICATION DE L'USAGER

Lorsque l'application est lancée, la première fenêtre (figure 10) à apparaître est celle où l'usager doit s'identifier. Deux types d'usager ont été définis, soit un usager dit expert, soit un second dit normal. En effet, les explications qui sont générées par la suite seront entre autre basées sur le profil de l'usager qui se retrouve dans la base des profils utilisateurs mentionnée dans l'architecture générale. De plus, il est possible d'accéder au prototype sous le titre d'administrateur.

Pour cette version du prototype, trois utilisateurs sont enregistrés sous les noms suivants : userN (usager normal), userE (usager expert) et admin (administrateur). Les mots de passe associés aux utilisateurs sont identiques aux noms d'usager, soit userN pour l'usager userN et ainsi de suite. Une fois le nom d'utilisateur entré suivi du mot de passe adéquat, le bouton *Enter* permet de passer à la fenêtre principale.

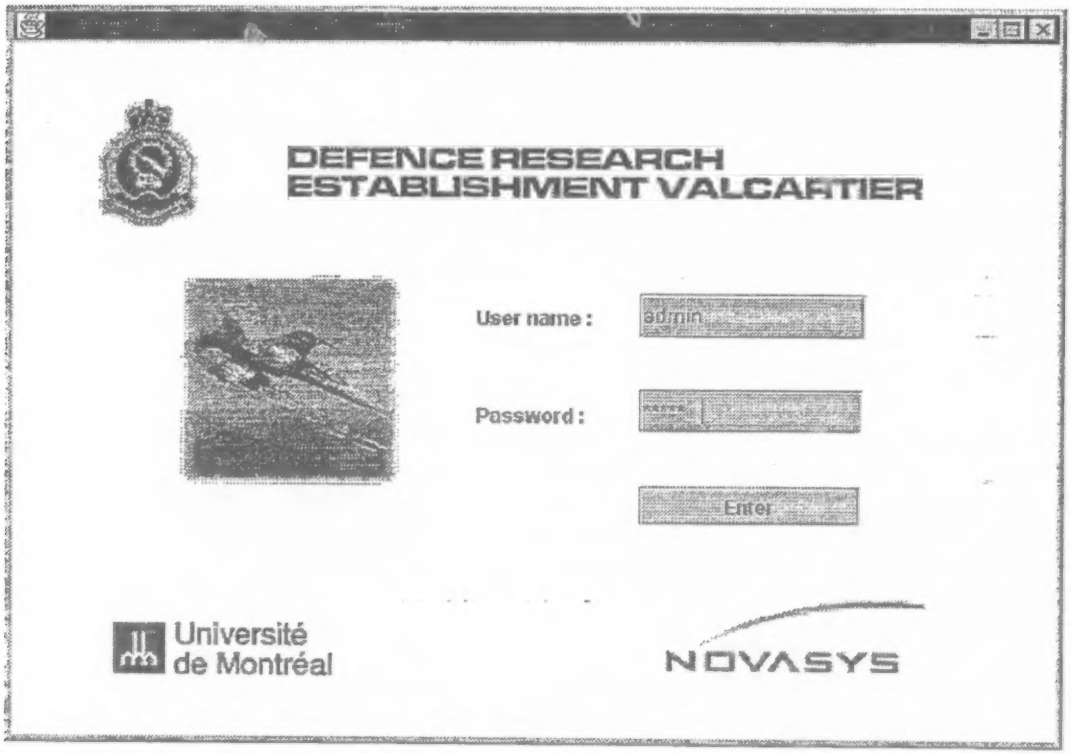

Figure 10 – Écran de présentation

7.4.3 MISE EN SITUATION

La fenêtre principale (figure 11) est divisée en trois parties distinctes. La première partie qui est composée de trois onglets intitulés *Situation*, *Parameters* et *Knowledge Base* est celle décrite dans cette section. L'onglet ayant comme titre *Situation* décrit ce que les différents senseurs ont capté, et ceci en langage naturel. Il est à noter que les mots soulignés de couleur bleu peuvent être sélectionnés afin d'avoir une définition plus détaillée. Ceci est directement relié à la base de connaissances terminologiques comportant la description de certains termes utilisés dans la génération d'explications.

L'onglet *Parameters* se veut une description plus en détails de la situation décrite sous le premier onglet. Initialement, le COA (course of action) est inconnu et pourra être précisé plus tard par le bouton *New Event* à la droite du tableau de paramètres (figure 12).

Le troisième onglet « *Knowledge Base* » est accessible seulement si l'usager possède le titre d'administrateur. Sous cet onglet, il est possible de visualiser toutes les informations pertinentes utilisées dans la génération de solutions et d'explications.

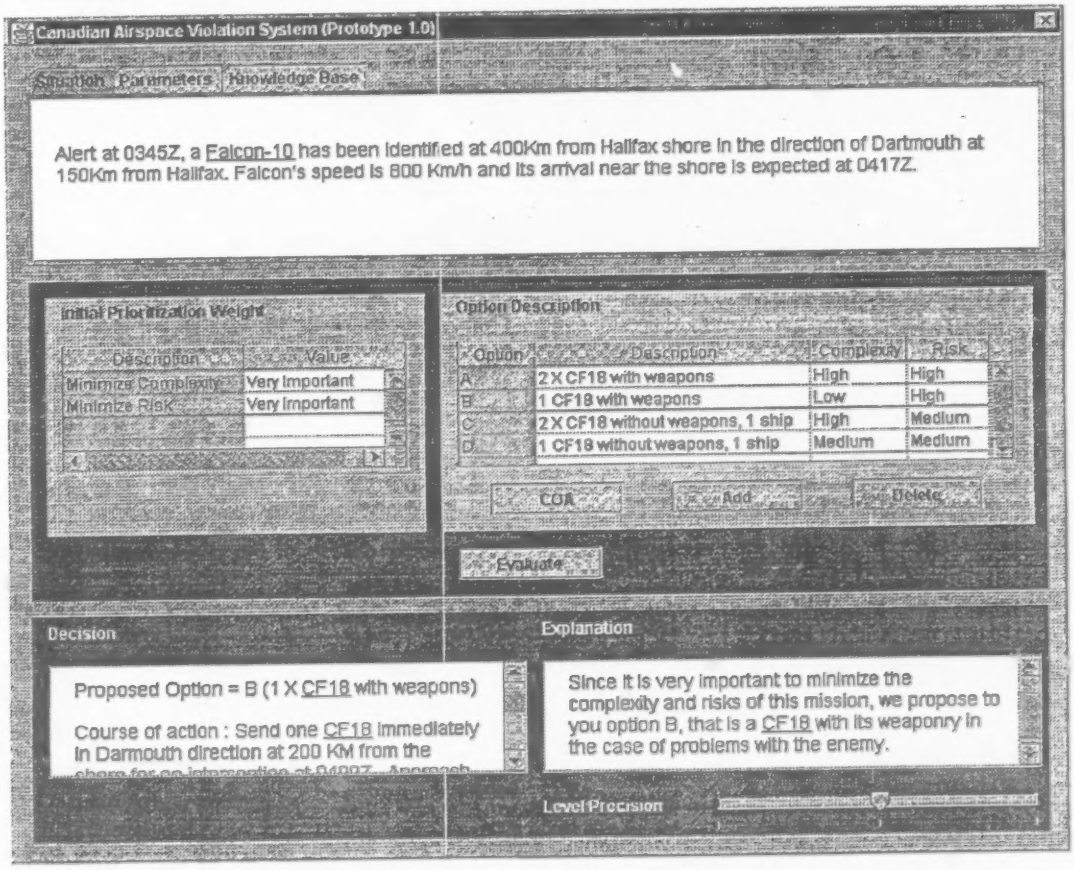

Figure 11 - Écran principal (onglet situation)

7.4.2 IDENTIFICATION DE L'USAGER

Lorsque l'application est lancée, la première fenêtre (figure 10) à apparaître est celle où l'usager doit s'identifier. Deux types d'usager ont été définis, soit un usager dit expert, soit un second dit normal. En effet, les explications qui sont générées par la suite seront entre autre basées sur le profil de l'usager qui se retrouve dans la base des profils utilisateurs mentionnée dans l'architecture générale. De plus, il est possible d'accéder au prototype sous le titre d'administrateur.

Pour cette version du prototype, trois utilisateurs sont enregistrés sous les noms suivants : userN (usager normal), userE (usager expert) et admin (administrateur). Les mots de passe associés aux utilisateurs sont identiques aux noms d'usager, soit userN pour l'usager userN et ainsi de suite. Une fois le nom d'utilisateur entré suivi du mot de passe adéquat, le bouton *Enter* permet de passer à la fenêtre principale.

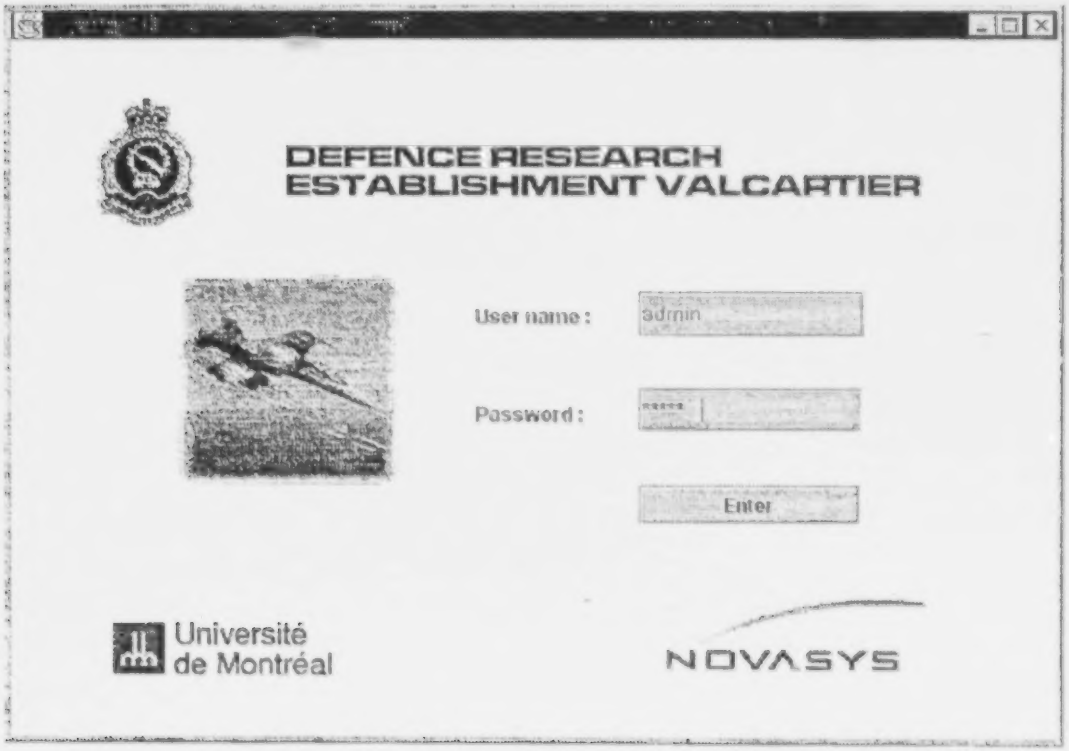

Figure 10 – Écran de présentation

7.4.3 MISE EN SITUATION

La fenêtre principale (figure 11) est divisée en trois parties distinctes. La première partie qui est composée de trois onglets intitulés *Situation*, *Parameters* et *Knowledge Base* est celle décrite dans cette section. L'onglet ayant comme titre *Situation* décrit ce que les différents senseurs ont capté, et ceci en langage naturel. Il est à noter que les mots soulignés de couleur bleu peuvent être sélectionnés afin d'avoir une définition plus détaillée. Ceci est directement relié à la base de connaissances terminologiques comportant la description de certains termes utilisés dans la génération d'explications.

L'onglet *Parameters* se veut une description plus en détails de la situation décrite sous le premier onglet. Initialement, le COA (course of action) est inconnu et pourra être précisé plus tard par le bouton *New Event* à la droite du tableau de paramètres (figure 12).

Le troisième onglet « *Knowledge Base* » est accessible seulement si l'usager possède le titre d'administrateur. Sous cet onglet, il est possible de visualiser toutes les informations pertinentes utilisées dans la génération de solutions et d'explications.

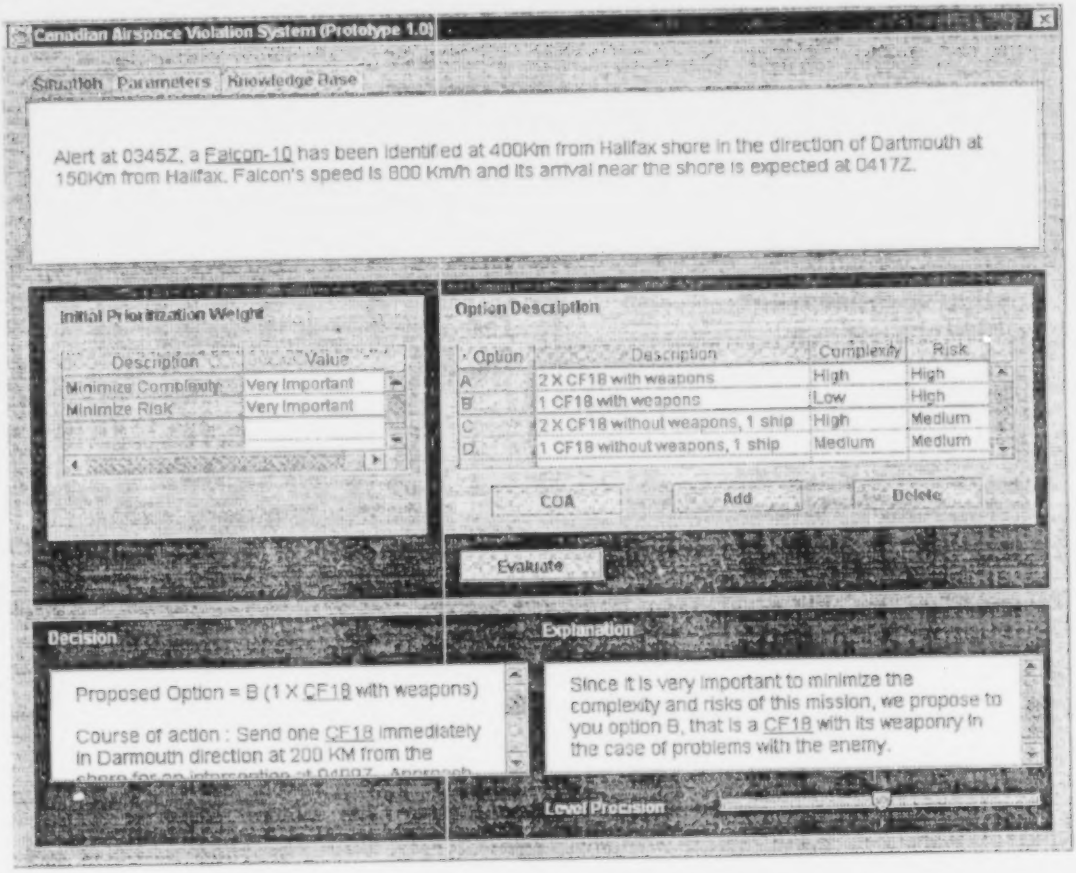

Figure 11 - Écran principal (onglet situation)

7.4.4 CHOIX DES POIDS ET OPTIONS

Le milieu de l'écran se divise en deux parties distinctes : du côté gauche, les différents paramètres associés à la future mission et du côté droit, les différentes options proposées à l'usager. Dans la partie gauche, le premier paramètre est la complexité, l'usager doit préciser quel est l'importance de minimiser la complexité. Il en est de même avec le second paramètre concernant l'importance de minimiser les risques lors de la mission. Une fois les poids choisis, le bouton *Evaluate* jusqu'ici inactif devient actif; il en est de même pour le bouton *New Event* situé dans l'onglet *Parameters* de la première section. Le *bouton New Event* ainsi actionné permet de mettre à jour le paramètre COA et ainsi d'avoir une décision valable suivie des explications.

Trois actions sont proposées à l'usager dans la section de droite. La première est la visualisation de l'option plus en détails par l'entremise du bouton *COA*. Les deuxième et troisième vont de paire puisqu'elles permettent d'ajouter ou d'éliminer une option. Ces ajouts ou suppressions devraient être prises en charge afin qu'un certain apprentissage soit fait en vue de futures situations.

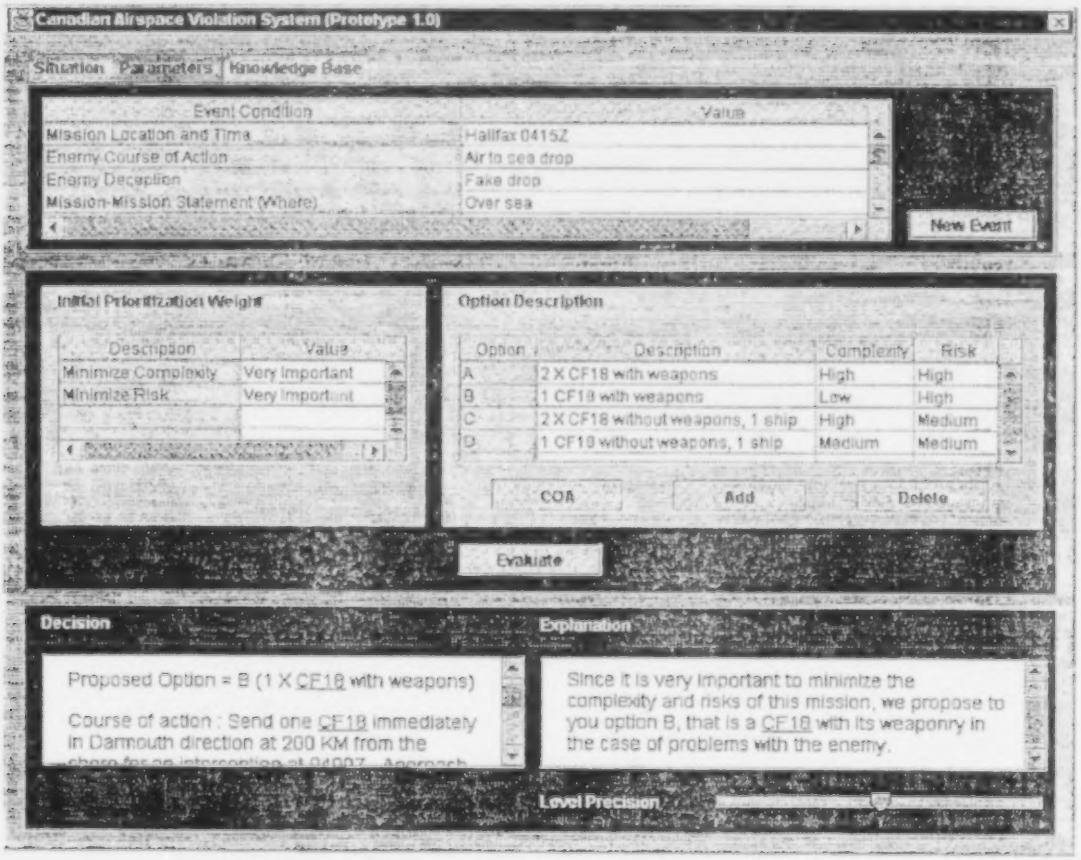

Figure 12 – Écran principal (onglet Parameters)

7.4.4 CHOIX DES POIDS ET OPTIONS

Le milieu de l'écran se divise en deux parties distinctes : du côté gauche, les différents paramètres associés à la future mission et du côté droit, les différentes options proposées à l'usager. Dans la partie gauche, le premier paramètre est la complexité, l'usager doit préciser quel est l'importance de minimiser la complexité. Il en est de même avec le second paramètre concernant l'importance de minimiser les risques lors de la mission. Une fois les poids choisis, le bouton *Evaluate* jusqu'ici inactif devient actif; il en est de même pour le bouton *New Event* situé dans l'onglet *Parameters* de la première section. Le *bouton New Event* ainsi actionné permet de mettre à jour le paramètre COA et ainsi d'avoir une décision valable suivie des explications.

Trois actions sont proposées à l'usager dans la section de droite. La première est la visualisation de l'option plus en détails par l'entremise du bouton *COA*. Les deuxième et troisième vont de paire puisqu'elles permettent d'ajouter ou d'éliminer une option. Ces ajouts ou suppressions devraient être prises en charge afin qu'un certain apprentissage soit fait en vue de futures situations.

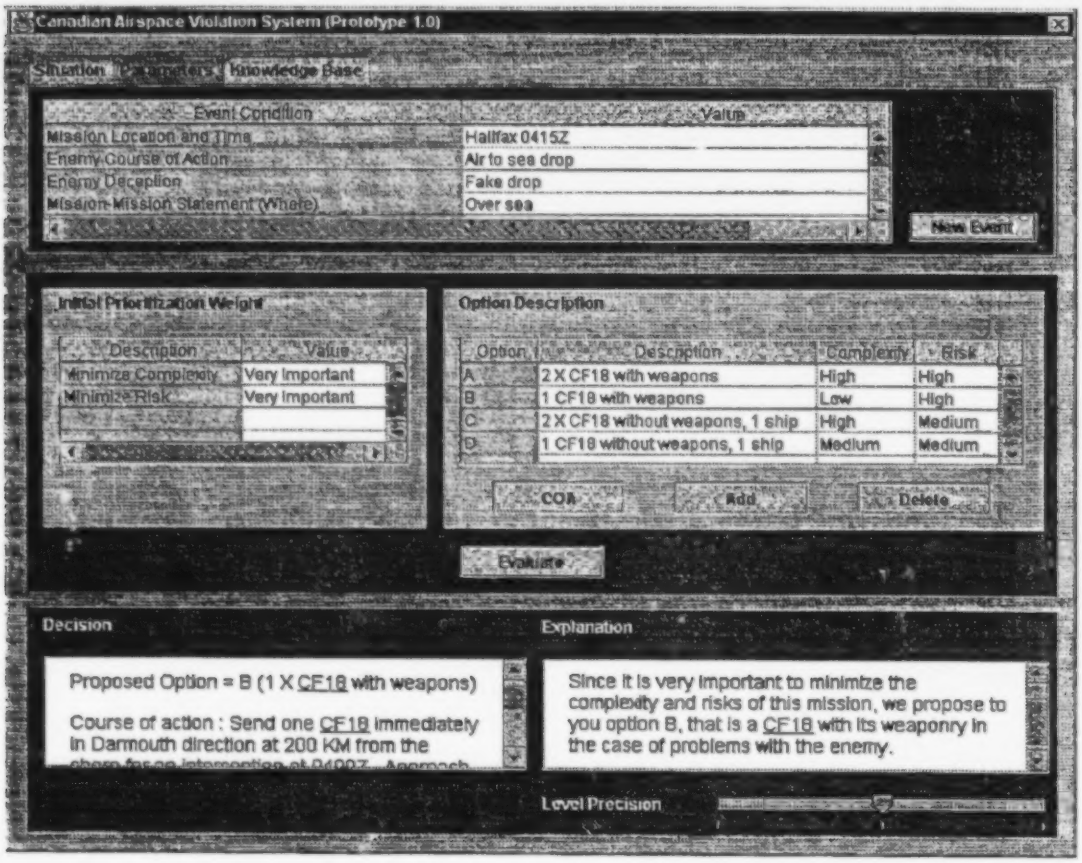

Figure 12 – Écran principal (onglet Parameters)

7.4.5 DÉCISION ET EXPLICATION

Cette dernière section se retrouve au bas de l'écran. Comme le titre l'indique, elle est divisée en deux parties, d'un côté sont présentées l'option et les actions (COA) proposées et de l'autre les explications qui y sont associées. De la même façon que dans la première section, les mots soulignés de couleur bleu peuvent être sélectionnés afin d'avoir une description du terme plus en détails. Au début inactive, cette section est activée par les boutons *Evaluate* ou *New Event*, le second apportant de nouvelles précisions sur la situation.

La section de droite portant sur les explications est accompagnée par une échelle de précision qui permet de générer différents niveaux d'explications en fonction du type d'usager. Ainsi, il est possible pour un usager normal de consulter les explications qui auraient été données à un usager expert ou vice versa.

7.5 PARAMÈTRES DU PROTOTYPE

Cette section présente les diverses options, COA, explication utilisée dans le développement de ce prototype.

7.5.1 MISE EN SITUATION

Alert at 0345Z, a Falcon-10 has been identified at 400Km from Halifax shore in the direction of Dartmouth at 150Km from Halifax. Falcon's speed is 800 Km/h and its arrival near the shore is expected at 0417Z.

7.5.2 PARAMÈTRES À 0345Z

Parameter	Value
Mission Location and Time: Where the mission is expected to be conducted.	Halifax 0415Z
Enemy Course of Action:	Not known
Enemy Deception:	fake drop
Mission-Mission Statement (Where):	over seas
Mission-Mission Statement (When):	as soon as possible
Mission-Mission Statement (When):	Before
Objectives:	obtain drug runners / obtain shipment / continue monitoring
Environment-Weather:	red
Forces Assigned: Friendly forces.	CF18 and ship
Enemy Forces-Aircraft and Weapon:	Dassault Falcon 10, Weapon not known
Enemy Forces-Personnel and Skills:	good
Enemy Forces-Communications Capabilities:	Radio / Morse Code Units

7.5.3 OPTIONS ET ACTIONS (COA) ASSOCIÉES

Option	Complexity Level	Risk Level	Solution	COA
A	High	High	2 CF18s with weapons	• Send the two CF18s immediately in Darmouth direction at 200 KM from the shore for an interception at 0400Z. • Approach the enemy's fighter plane with one of the two CF18s in order to identify it and check its weaponry. • The other CF should stay close to the enemy's fighter plane and within its vision scope in order to show our strike power. • Escort the enemy's fighter plane to Halifax Airport in order to verify its cargo and do the usual interrogation. • In the case of invasion or attack, contact the commander office and follow the usual procedures.
B	Low	High	1 CF18 with weapons	• Send one CF18 immediately in Darmouth direction at 200 KM from the shore for an interception at 0400Z. • Approach the enemy's fighter plane with the CF18 in order to identify it and check its weaponry. • Escort the enemy's fighter plane to Halifax Airport in order to verify its cargo and do the usual interrogation. • In the case of invasion or attack, contact the commander office and follow the usual procedures.

7.5 PARAMÈTRES DU PROTOTYPE

Cette section présente les diverses options, COA, explication utilisée dans le développement de ce prototype.

7.5.1 MISE EN SITUATION

Alert at 0345Z, a Falcon-10 has been identified at 400Km from Halifax shore in the direction of Dartmouth at 150Km from Halifax. Falcon's speed is 800 Km/h and its arrival near the shore is expected at 0417Z.

7.5.2 PARAMÈTRES À 0345Z

Parameter	Value
Mission Location and Time: Where the mission is expected to be conducted.	Halifax 0415Z
Enemy Course of Action:	Not known
Enemy Deception:	fake drop
Mission-Mission Statement (Where):	over seas
Mission-Mission Statement (When):	as soon as possible
Mission-Mission Statement (When):	Before
Objectives:	obtain drug runners / obtain shipment / continue monitoring
Environment-Weather:	red
Forces Assigned: Friendly forces.	CF18 and ship
Enemy Forces-Aircraft and Weapon:	Dassault Falcon 10, Weapon not known
Enemy Forces-Personnel and Skills:	good
Enemy Forces-Communications Capabilities:	Radio / Morse Code Units

7.5.3 OPTIONS ET ACTIONS (COA) ASSOCIÉES

Option	Complexity Level	Risk Level	Solution	COA
A	High	High	2 CF18s with weapons	• Send the two CF18s immediately in Darmouth direction at 200 KM from the shore for an interception at 0400Z. • Approach the enemy's fighter plane with one of the two CF18s in order to identify it and check its weaponry. • The other CF should stay close to the enemy's fighter plane and within its vision scope in order to show our strike power. • Escort the enemy's fighter plane to Halifax Airport in order to verify its cargo and do the usual interrogation. • In the case of invasion or attack, contact the commander office and follow the usual procedures.
B	Low	High	1 CF18 with weapons	• Send one CF18 immediately in Darmouth direction at 200 KM from the shore for an interception at 0400Z. • Approach the enemy's fighter plane with the CF18 in order to identify it and check its weaponry. • Escort the enemy's fighter plane to Halifax Airport in order to verify its cargo and do the usual interrogation. • In the case of invasion or attack, contact the commander office and follow the usual procedures.

Option	Complexity Level	Risk Level	Solution	COA
C	High	Medium	2 CF18s without weapons + 1 ship	• Send the two CF18s immediately in Darmouth direction at 200 KM from the shore for an interception at 0400Z. • Approach with precaution the enemy's fighter plane with one of the two CF18s in order to check its weaponry and identify it. • The other CF18 should stay close to the enemy's fighter plane and within its vision scope in order to show our pacific intervention. • Escort the enemy's fighter plane to Halifax Airport in order to verify its cargo and do the usual interrogation. • Make sure that no object was released by the enemy's plane in the sea. • In the case of invasion or attack, contact the commander office and follow the usual procedures.
D	Medium	Medium	1 CF18 without weapons + 1 ship	• Send one CF18 immediately in Darmouth direction at 200 KM from the shore for an interception at 0400Z. • Send the ship in order to make sure that no object was released in the sea by the enemy's fighter plane. • Approach the enemy's fighter plane with the CF18 in order to identify it and check its weaponry. • Escort the enemy's fighter plane to Halifax Airport in order to verify its cargo and do the usual interrogation. • In the case of invasion or attack, contact the commander office and follow the usual procedures.

7.5.4 EXPLICATIONS SELON LES OBJECTIFS ET LES OPTIONS

7.5.4.1 Enemy COA = Testing run

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Very important	Very important	B	Since it is very important to minimize the complexity and risks of this mission, we propose to you option B, that is a CF18 with its weaponry in the case of problems with the enemy.	Since it is very important to minimize the complexity of this mission, we propose to you an option that doesn't imply sending the ship and a minimum of high speed fighter planes. The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is very important to minimize the risks of this mission, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option B.

Option	Complexity Level	Risk Level	Solution	COA
C	High	Medium	2 CF18s without weapons + 1 ship	• Send the two CF18s immediately in Darmouth direction at 200 KM from the shore for an interception at 0400Z. • Approach with precaution the enemy's fighter plane with one of the two CF18s in order to check its weaponry and identify it. • The other CF18 should stay close to the enemy's fighter plane and within its vision scope in order to show our pacific intervention. • Escort the enemy's fighter plane to Halifax Airport in order to verify its cargo and do the usual interrogation. • Make sure that no object was released by the enemy's plane in the sea. • In the case of invasion or attack, contact the commander office and follow the usual procedures.
D	Medium	Medium	1 CF18 without weapons + 1 ship	• Send one CF18 immediately in Darmouth direction at 200 KM from the shore for an interception at 0400Z. • Send the ship in order to make sure that no object was released in the sea by the enemy's fighter plane. • Approach the enemy's fighter plane with the CF18 in order to identify it and check its weaponry. • Escort the enemy's fighter plane to Halifax Airport in order to verify its cargo and do the usual interrogation. • In the case of invasion or attack, contact the commander office and follow the usual procedures.

7.5.4 EXPLICATIONS SELON LES OBJECTIFS ET LES OPTIONS

7.5.4.1 Enemy COA = Testing run

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Very important	Very important	B	Since it is very important to minimize the complexity and risks of this mission, we propose to you option B, that is a CF18 with its weaponry in the case of problems with the enemy.	Since it is very important to minimize the complexity of this mission, we propose to you an option that doesn't imply sending the ship and a minimum of high speed fighter planes. The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is very important to minimize the risks of this mission, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option B.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Very important	Fairly important	B	Since it is very important to minimize the complexity of this mission by taking the risk into consideration, we propose to you option B, that is a CF18 with its weaponry in the case of problems with the enemy.	Since it is very important to minimize the complexity of this mission, we propose to you an option that doesn't imply sending the ship and a minimum of high speed fighter planes. The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since the risks of this mission should be taken into consideration, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option B.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Very important	Not important	B	Since it is very important to minimize the complexity of this mission without worrying about the risks, we propose to you option B, that is, a CF18 which implies a minimum of high speed fighter planes.	Since it is very important to minimize the complexity of this mission, we propose to you an option that doesn't imply sending the ship and a minimum of high speed fighter planes. The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In that case, we recommend to you option B.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Very important	Fairly important	B	Since it is very important to minimize the complexity of this mission by taking the risk into consideration, we propose to you option B, that is a CF18 with its weaponry in the case of problems with the enemy.	Since it is very important to minimize the complexity of this mission, we propose to you an option that doesn't imply sending the ship and a minimum of high speed fighter planes. The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since the risks of this mission should be taken into consideration, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option B.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Very important	Not important	B	Since it is very important to minimize the complexity of this mission without worrying about the risks, we propose to you option B, that is, a CF18 which implies a minimum of high speed fighter planes.	Since it is very important to minimize the complexity of this mission, we propose to you an option that doesn't imply sending the ship and a minimum of high speed fighter planes. The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In that case, we recommend to you option B.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Fairly important	Very important	B	Since it is very important to minimize the risk of this mission by taking the complexity into consideration, we propose to you option B, that is a CF18 with its weaponry in the case of problems with the enemy.	Since the complexity of this mission should be taken into consideration, we propose to you an option that doesn't imply sending a ship and a minimum of high speed fighter planes. The proposed fighter plane is a CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is very important to minimize the risks of this mission, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option B.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Fairly important	Fairly important	D	Since it is important to take into consideration the complexity and the risks of the mission, we propose to you option D, that is, a CF18 without its weaponry and a ship.	Since the complexity of this mission should be taken into consideration, we propose to you an option that implies sending a minimum of ships and a minimum of high speed fighter planes. The proposed fighter plane is a CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since the risks of this mission should be taken into consideration, we recommend that you send a marine unit combined with an airforce unit. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, is could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option D.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Fairly important	Not important	D	Since it is important to take into consideration the complexity of the mission without worrying about the risk, we propose to you option D, that is a CF18 without its weaponry and a ship.	Since it is important to take into consideration the complexity of this mission, we propose to you an option that doesn't imply sending the ship and a minimum of high speed fighter planes. The proposed fighter plane is the CF18, it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In that case, we recommend to you option D.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Not important	Very important	A	Since it is very important to minimize the risks of this mission without worrying about the risk, we propose to you option A, that is two CF18s with their weaponry in case of problems with the enemy.	Since it is not important to minimize the complexity of this mission, we propose to you an option that implies sending two high speed fighter planes. The proposed fighter planes are CF18s, they answer the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is very important to minimize the risks of this mission, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option A.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Not important	Fairly important	C	Since it is important to take into consideration the risk of the mission without worrying about the complexity, we propose to you option C, that is two CF18s without its weaponry and a ship.	Since it is not important to minimize the complexity of this mission, we propose to you an option that implies sending one ship and two high speed fighters planes. The proposed fighters planes are two CF18s, since they answer the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is important to take into sonsideration the risks of this mission, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option A.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Not important	Not important	D	Since it doesn't matter much to minimize the complexity or the risks of the mission, we propose to you option D, that is a CF18 without its weaponry and a ship.	Since it is not important to minimize the complexity of this mission, we propose to you an option that implies sending one ship and one high speed fighter planes. The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is not important to minimize the risks of this mission, we recommend that you not arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option D.

7.5.4.2 Enemy COA = Air to sea drop

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Very important	Very important	B	Since it is very important to minimize the complexity and risks of this mission, we propose to you option B, that is a CF18 with its weaponry in the case of problems with the enemy.	Since it is very important to minimize the complexity of this mission, we propose to you an option that doesn't imply sending the ship and a minimum of high speed fighter planes. The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is very important to minimize the risks of this mission, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option B.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Very important	Fairly important	D	Since it is very important to minimize the complexity of this mission by taking the risk into consideration, we propose to you option D, that is a CF18 without its weaponry and a ship.	Since it is very important to minimize the complexity of this mission, we propose to you an option that imply sending at least one ship and a minimum of high speed fighter planes.The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since the risks of this mission should be taken into consideration, we recommend that you send a marine unit combined with an airforce unit. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option D.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Very important	Not important	D	Since it is very important to minimize the complexity of the mission without worrying about the complexity, we propose to you option D, that is a CF18 without its weaponry and a ship.	Since it is very important to minimize the complexity of this mission, we propose to you an option that implies sending at least one ship and a minimum of high speed fighter planes.The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is not important to minimize the risks of this mission, we recommend that you not arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss; this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option D.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Fairly important	Very important	A	Since it is very important to minimize the risk of this mission by taking the complexity into consideration, we propose to you option A, that is two CF18 with their weaponry in the case of problems with the enemy.	Since the complexity of this mission should be taken into consideration, we propose to you an option that implies sending two high speed fighter planes. The proposed fighter planes are two CF18s, since they answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is very important to minimize the risks of this mission, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option A.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Fairly important	Fairly important	D	Since it is important to take into consideration the complexity and the risks of the mission, we propose to you option D, that is, a CF18 without its weaponry and a ship.	Since the complexity of this mission should be taken into consideration, we propose to you an option that implies sending a minimum of ships and a minimum of high speed fighter planes. The proposed fighter plane is a CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since the risks of this mission should be taken into consideration, we recommend that you send a marine unit combined with an airforce unit. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option D.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Fairly important	Not important	D	Since it is important to take into consideration the complexity of the mission without worrying about the risk, we propose to you option D, that is a CF18 without its weaponry and a ship.	Since it is important to take into consideration the complexity of this mission, we propose to you an option that doesn't imply sending the ship and a minimum of high speed fighter planes.The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In that case, we recommend to you option D.
Not important	Very important	A	Since it is very important to minimize the risks of this mission without worrying about the risk, we propose to you option A, that is two CF18s with their weaponry in case of problems with the enemy.	Since it is not important to minimize the complexity of this mission, we propose to you an option that doesn't imply sending two high speed fighter planes. The proposed fighter plane is a CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is very important to minimize the risks of this mission, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option A.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Not important	Fairly important	C	Since it is important to take into consideration the risk of the mission without worrying about the complexity, we propose to you option C, that is two CF18s without their weaponry and a ship.	Since it is not important to minimize the complexity of this mission, we propose to you an option that implies sending one ship and two high speed fighter planes. The proposed fighter planes are two CF18s, since they answer the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is important to take into consideration the risks of this mission, we recommend that you arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option A.

Min. Comp.	Min. Risk	Option	Explanation Expert	Explanation Normal
Not important	Not important	D	Since it doesn't matter much to minimize the complexity or the risks of the mission, we propose to you option D, that is a CF18 without its weaponry and a ship.	Since it is not important to minimize the complexity of this mission, we propose to you an option that implies sending one ship and one high speed fighter planes. The proposed fighter plane is the CF18, since it answers the need for speed required for interception. The complexity is concerned with how difficult the implementation of a COA would be. In addition, since it is not important to minimize the risks of this mission, we recommend that you not arm yourselves. The risk is concerned with the possibility of mission failure and loss of equipment. It should be noted that bad weather potentially impacts the following risks and should be considered when evaluating the following risks. In addition, while a confrontation could cause personnel loss, this could be due to other causes. Thus personnel loss and confrontation risk should be considered as separate items. In that case, we recommend to you option D.

8. CONCLUSION

Lors de cette revue bibliographique, nous avons survolé les principales méthodes pouvant améliorer les capacités explicatives des systèmes d'aide à la décision. Nous avons présenté les fondements de base des explications pour ensuite présenter ces différentes méthodes. Le premier point qui se dégage de cette étude est que les mécanismes de génération d'explications doivent être considérés comme un problème et un défi à part entière. En effet, expliquer un raisonnement est une tâche de résolution de problèmes assez complexe que nous pouvons considérer comme étant une tâche intelligente. Pour expliquer son raisonnement, un système doit examiner toutes les connaissances mises en jeu pour aboutir à une solution et comprendre les raisonnements mis en œuvre pour résoudre le problème. Autrement dit, pour expliquer son raisonnement, un système doit raisonner sur la manière dont il résout les problèmes.

Un second point qui nous semble important est le fait que les explications générées doivent être construites en fonction du niveau de connaissances de l'utilisateur. Trop souvent, les explications générées sont produites et pensées pour les experts du domaine d'application. Il est alors impossible d'établir un niveau suffisant de confiance entre l'utilisateur et le système d'aide à la décision.

La conception d'un SBR ou plus généralement d'un système basé sur la connaissance soulève plusieurs questions : quelles connaissances fournir au système, comment les représenter, quels mécanismes définir pour accéder à ces connaissances et pour les interpréter? Lorsqu'il s'agit de système d'aide à la décision, on doit étudier les mêmes questions, mais de deux points de vue différents : la résolution de problèmes et l'explication du raisonnement. Autrement dit, on doit permettre d'utiliser ces sources de connaissances et les connaissances qui permettent d'utiliser ces sources en vue d'accomplir la double tâche du système. D'autre part, comme nous l'avons dit plus haut, nous considérons l'explication comme une tâche de résolution de problèmes.

À travers les diverses techniques explorées dans cette recherche, plusieurs points communs sont mis en évidence. En premier lieu, que ce soit par une base de connaissances spécifique au profil de l'utilisateur ou par plusieurs agents intelligents ayant une granularité différente, le type d'utilisateur du système est un facteur majeur sur le type et la forme des explications. De plus, la plupart des modèles possèdent une base de connaissances dédiée à la génération d'explications. Ces connaissances peuvent être gérées par des méta-connaissances ayant un niveau d'abstraction plus élevé. Il est ainsi possible de visualiser la stratégie globale du système et non seulement des explications propres à un fait en particulier. Le dernier point abordé qui est celui des agents intelligents pouvant générer des explications ou expliquer leur propre comportement lors de simulations spécifiques est une approche nouvelle qui semble très prometteuse.

9. PERSONNES CONTACTS DU DOMAINE

PAYS	CONTACTS	ADRESSES
Canada	Brian Gaines	Knowledge Science Institute University of Calgary Calgary T2N 1N4 Alberta Canada E-mail: gaines@cpsc.ucalgary.ca Web : http://www.cpsc.ucalgary.ca/projects/grouplab/699/gaines.html
Canada	Nicole Tourigny	Département d'informatique Local 3721, Pavillon Adrien-Pouliot Université Laval Sainte-Foy, Canada G1K 7P4 E-mail: Nicole.Tourigny@IFT.ulaval.ca Web : http://employes.ift.ulaval.ca/~tourigny/
France	Patrick Brezillon	LIP6, Box 169, University Paris VI, 4 Place Jussieu, 75252 PARIS Cedex 05, France Phone: +33 1 44 27 70 08 - Fax: +33 1 44 27 70 00 E-mail: Patrick.Brezillon@lip6.fr Web: http://www-sysdef.lip6.fr/~brezil
France	Michelle Joab	Laboratoire LIP6, Université P. et M.Curie, Tour 55-65, bureau 330 4, place jussieu, 75 252 Paris cedex 05 Tel: 01 44 27 71 48 Fax: 01 44 27 70 00 E-mail: Michelle.Joab@lip6.fr Web : http://www-sysdef.lip6.fr/SYSDEF/membres/permanents/Michele-Joab.html
France	Benoit Lemaire et Brigitte Safar	E-mail : blemaire@brassens.upmf-grenoble.fr (Le groupe est actuellement dissout)

9. PERSONNES CONTACTS DU DOMAINE

PAYS	CONTACTS	ADRESSES
Canada	Brian Gaines	Knowledge Science Institute University of Calgary Calgary T2N 1N4 Alberta Canada E-mail: gaines@cpsc.ucalgary.ca Web : http://www.cpsc.ucalgary.ca/projects/grouplab/699/gaines.html
Canada	Nicole Tourigny	Département d'informatique Local 3721, Pavillon Adrien-Pouliot Université Laval Sainte-Foy, Canada G1K 7P4 E-mail: Nicole.Tourigny@IFT.ulaval.ca Web : http://employes.ift.ulaval.ca/~tourigny/
France	Patrick Brezillon	LIP6, Box 169, University Paris VI, 4 Place Jussieu, 75252 PARIS Cedex 05, France Phone: +33 1 44 27 70 08 - Fax: +33 1 44 27 70 00 E-mail: Patrick.Brezillon@lip6.fr Web: http://www-sysdef.lip6.fr/~brezil
France	Michelle Joab	Laboratoire LIP6, Université P. et M.Curie, Tour 55-65, bureau 330 4, place jussieu, 75 252 Paris cedex 05 Tel: 01 44 27 71 48 Fax: 01 44 27 70 00 E-mail: Michelle.Joab@lip6.fr Web : http://www-sysdef.lip6.fr/SYSDEF/membres/permanents/Michele-Joab.html
France	Benoit Lemaire et Brigitte Safar	E-mail : blemaire@brassens.upmf-grenoble.fr (Le groupe est actuellement dissout)

PAYS	CONTACTS	ADRESSES
France	Unité de recherche INRIA Sophia Antipolis	2004, route des Lucioles – B.P. 93 06902 Sophia Antipolis Cedex France Téléphone : 04 92 38 77 77 Télécopie : 04 92 38 77 65 **En particulier le projet** ACACIA Acquisition des connaissances pour l'assistance à la conception par interaction entre agents **Responsable** : Rose DIENG +33 4 92 38 78 10 Rose.Dieng@inria.fr Web : http://www.inria.fr/Unites/Sophia-fra.html
États-Unis	Christine M. Mitchell	Center for Human-Machine Systems Research School of Industrial and Systems Engineering 765 Ferst St. Georgia Institute of Technology Atlanta, GA 30332-0205 USA +1 404-894-4321 +1 404-385-0357 E-mail : cm@chmsr.gatech.edu Web: http://www.isye.gatech.edu/chmsr/Christine_Mitchell

10. BIBLIOGRAPHIE

[Breuker & *Wielinga*, 89] Breuker, J., Wielinga, B. **Models of expertise in knowledge acquisition**. *In Topics in expert system design:methodologies and tools*, Guida, G., Tasso, C., editors, North Holland Publishing Company, Amsterdam, 1989.

[Brown et Chandrasekaran, 89] Brown, D., Chandrasekaran, B (1989). **Design Problem Solving: Knowledge** Structures and Control Strategies. London: Pitman.

[Chandrasekaran *et al*, 88] Chandrasekaran, B., Tanner, M.C., Josephson, J.R. **Explanation: the role of control strategies and deep models**. *In Expert systems: the user interface*, Ablex Publishing Corporation, 1988.

[Clancey, 83] Clancey, W.J. **The Epistemology of a Rule-Based Expert System: A Framework for Explanations**. *Stanford Heuristic Programming Project*, Report HPP-83-17, 1983.

[Clancey, 86] Clancey, W.J. **From GUIDON to NEOMYCIN and HERACLES in Twenty Short Lessons: ORN Final Report 1979-1985**. AI Magazine, 7(3), pp 40-60, 1986.

[Clancey, 87] Clancey, W.J., **Knowledge-Based Tutoring – The GUIDON program**. The MIT Press, Cambridge, 1987.

[Couteret, 98] Couteret, Paul, **La place de la confiance dans la mobilisation des ressources humaines dans les petites entreprises : le cas de la relation dirigeant-adjoint. CIFPME, 1998.**

[Davis, 76] Davis, R. **Applications of meta-level knowledge to the construction, maintenance, and use of large knowledge bases**. Memo 283, Stanford University. Réimprimé en R. Davis et D. Lenat, Knowledge-Based Systems in Artificial Intelligence. McGraw Hill, New York, 1980.

[Dominguez, 90] Dominguez, C.J. **Sur l'explication dans les systèmes à base de règles : le système prose**. Thèse d'Université, Université Paris VI, pp 235, 1990.

[Findler, 79] Findler, N.V. **Associative Networks. Representation and Use of Knowledge by Computers**. Academic Press, New York, 1979.

[Frasson, 98] Frasson, C. **Using Cognitive Agents for Building Pedagogical Strategies in a Multistrategic Intelligent Tutoring System**, *Deuxième journée Acteurs, Agents et Apprentissage*, Bayonne, September 1998.

[Goguen *et al*, 83] Goguen, J.A., Weiner, J.L., Linde, C. **Reasoning and natural explanation**. International Journal Man-Machine Studies, 19, 1983.

[Gréboval, 96] Gréboval, M.H. **Vers une représentation du dialogue explicatif**. *In Actes de la conférence : Explication'96*, France, pp 263-274, 1996.

[Guindon, 90] Guindon, R. **Knowledge exploited by experts during software system design**. *International Journal of Man-Machine Studies*, 33, pp 279-304, 1990.

[Haton *et al*, 91] Haton, J.P, Bouzid, N., Charpillet, F. Haton, M.C, Laasri, B., Laasri, H., Marquis, P., Mondot, T. and Napoli, A. **Le raisonnement en intelligence artificielle**. InterEditions, Paris, France, pp 480, 1991.

[Hayes-Roth, 95] Hayes-Roth, B., "**An architecture for adaptive intelligent systems**". *Artificial Intelligence: special issue on agents and interactivity*, (pp. 327-365), 1995.

[Herrmann et al., 98] Herrmann, Jürgen, Kloth, Matthias, Feldkamp, Frank, "**The Role of Explanations in an Intelligent Assistant System**". *Artificial Intelligence in engineering, Elsevier, vol 12, (pp. 107-126), 1998.*

[Johnson, 94] Johnson, W.L. **Agents that Learn to Explain Themselves**. *In proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI'94)*, Seatle, Etats-Unis, pp 1257-1263, 1994.

[Kassel, 89] Kassel, G. **Systèmes Experts de Seconde Génération & Explication du Raisonnement**, *In Actes de la conférence : Les systèmes experts et leurs applications*, Avignon, France, pp 223-243, 1989.

[Karbach *et al*, 91] Karback, W., Vob, A., Schuckey, R., Drouven, U. **MODEL-K: prototyping at the knowledge level**. In actes du congres KMET'91, pp 195-208, 1991.

[Kodratoff, 86] Kodratoff, Y. **Leçon d'Apprentissage Symbolique Automatique**. *Collection Techniques Avancées de l'Informatique*, Cepadués Editions, 1986.

[Krivine et David, 91] Krivine, J.P., David, J.M. **L'acquisition des connaissances vue comme un processus de modélisation : méthodes et outils**. *Intellictica*, 12, pp 101-137, 1991.

[Lazure, 88] Lazure, Noêl. **Vocabulaire de l'intelligence artificielle : Bulletin de terminologie 184.** *Ministère de l'approvisionnement et services du Canada*.p. 218, 1988

[Lemaire *et al*, 96] Lemaire, B., Safar, B. and Yonnet, C. **L'acquisition des connaissances d'explication**. In Aussenac-Gilles, N., Laublet, P. and Reynaud, C., editors, *Acquisition et ingénierie des connaissances : tendances actuelles.* Cépaduès éditions, France, pp 567-384, 1996.

[Levy-Leblond, 91] Levy-Leblond, J.M. **La physique, une science sans complexe?**, *In Les théoires de la complexité*, Seuil, pp 127-134, 1991.

[Nii, 86] Nii, H.P. **Blackboard systems: blackboard application systems, blackboard systems from a knowledge engineering perspective**, *AI Magazine*, 7(3), pp 82-106, 1986.

[Pitrat, 96] Pitrat, J. **Méta-expliquer pour apprendre une stratégie**. In Actes de la conférence : *Explication'96*, France, pp 123-134, 1996.

[Richard et al.] Richard, Jean-François et al.. **Le traitement de l'information symbolique**. Traité de psychologie cognitive 2. Éd. Dunod, 1992.

[Rolston, 88] Rolston. D.W. **Principles of Artificial Intelligence and Expert Systems Development**. McGraw-Hill, NewYork, pp 257, 1988.

[Schank, 82] Schank, R.C. **Explanation Patterns**. Editions Lawrence Erlbaum Associates, 1982.

[Schank et Leake, 89] Schank, R.C., Leake, D.B. **Creativity and Learning in a Case-Based Explainer**. *Artificial Intelligence an International Journal, Volume 40 :1-3*, pp 353-385, 1989.

[Shortliffe *et al*, 75] Shortliffe, E.H., Davis, R., Axline, S.G., Buchanan, B.G., Green, C.C., Cohen, S.N. **Computer-based consultation and clinical therapeutics: Explanation and rule acquisition capabilities of the MYCIN system**. Computers and Biomedcal Researsh 8, 1975.

[Shortliffe, 76] Shortliffe, E.H. **Computer-Based Medical Consultations: MYCIN**, Elsevier, New York, 1976.

[Swartout, 81] Swartout, W.R. **Producing explanations and justifications of expert consulting programs**. MIT, Laboratory for Computer Science, Rapport TR-251, 1981.

[Swartout et Moore, 93] Swartout, W.R. and Moore, J. **Explanation in Second Generation Expert Systems.** In David, J.M., Krivine, J.P. and Simmons, R., editors, *Second generation expert systems*. Springer Verlag, Berlin-New York, pp 543-585, 1993.

[Tadié *et al*, 96] Tadié, S., Frasson, C., Lefevre, B. **Coopération entre agents conseillers pour la génération d'explications dans un environnement coopératif**. *In Actes de la conférence : Explication'96*, France, pp 49-61, 1996.

[Tadié *et al*, 98] Tadié, S., Frasson, C., Lefebvre, B. **Un modèle de représentation de tâches coopératives permettant la génération d'explications dans les environnements distribués d'enseignement : MONACO_T**, *NTICF'98, Conférence internationale sur les nouvelles technologies de la communication et de la formation*, Rouen, November 1998.

[Tanner *et al*, 93] Tanner, M.C., Keuneke, A.M. and Chandrasekaran, B. **Explanation Using Task Structure and Domain Functional Models.** In David, J.M., Krivine, J.P. and Simmons, R., editors, *Second generation expert systems*. Springer Verlag, Berlin-NewYork, pp 586-613, 1993.

[Tecuci, 98] Tecuci, G., **Building Intelligent Agents: An apprenticeship Multistrategy Learning Theory Methodology, Tool and Cases Studies**, Academic Press, London, 1998.

[Waterman, 86] Waterman, Donald A. A Guide to Expert Systems. Ed. Addison-Wesley, 1986.

[Wick, 93] Wick, M.R. **Second Generation Expert System Explanation.** In David, J.M., Krivine, J.P. and Simmons, R., editors, *Second generation expert systems*. Springer Verlag, Berlin-NewYork, pp 614-640, 1993.

[Wick et Thompson, 92] Wick, M.R., Thompson, W.B. **Reconstructive expert system explanation**. *Artificial Intelligence an International Journal, Volume 50*, pp 33-70, 1992.

#511364